PB-245 009

STATUS OF NADB DATA SYSTEMS

David W. Armentrout, et al

PEDCo-Environmental Specialists, Incorporated
Cincinnati, Ohio

April 1975

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

268268

EPA-450/3-75-065

April 1975

PB 245 009

STATUS OF NADB DATA SYSTEMS

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
US Department of Commerce
Springfield, VA. 22151

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, North Carolina 27711

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.
EPA-450/3-75-065

2.

3. RECIPIENT'S ACCESSION NO.

4. TITLE AND SUBTITLE
Status of NADB Data Systems

5. REPORT DATE
April, 1975

6. PERFORMING ORGANIZATION CODE

7. AUTHOR(S)
David W. Armentrout
Charles E. Zimmer

8. PERFORMING ORGANIZATION REPORT NO.

9. PERFORMING ORGANIZATION NAME AND ADDRESS
PEDCo-Environmental Specialists, Inc.
Suite 13, Atkinson Square
Cincinnati, Ohio 45246

10. PROGRAM ELEMENT NO.

11. CONTRACT/GRANT NO.
68-02-1375
Task Order No. 10

12. SPONSORING AGENCY NAME AND ADDRESS
Environmental Protection Agency
Office of Air & Water Programs
National Air Data Branch
Research Triangle Park, North Carolina 27711

13. TYPE OF REPORT AND PERIOD COVERED
Final

14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This report presents the results of a survey conducted among the Regional Offices of the U. S. Environmental Protection Agency and other EPA users of the NADB systems. Included are recommendations for improving the current NADB systems to make them more responsive to users and detailed requirements for new programs are outlined. Costs for implementating each suggestion are estimated.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Data Retrieval Air Pollution	NADB NEDS SAROAD Air Quality Data	13B

PRICES SUBJECT TO CHANGE

18. DISTRIBUTION STATEMENT
Unlimited

19. SECURITY CLASS *(This Report)*
Unclassified

20. SECURITY CLASS *(This page)*
Unclassified

21. NO. OF PAGES
205

22. PRICE
$7.25

EPA Form 2220-1 (9-73)

EPA-450/3-75/065

STATUS OF NADB DATA SYSTEMS

by

PEDCo-Environmental Specialists, Inc.
Suite 13
Atkinson Square
Cincinnati, Ohio 45246

Contract No. 68-02-1375
Task No. 10

EPA Project Officer: Gerald Nehls

Prepared for

ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Waste Management
Office of Air Quality Planning and Standards
Research Triangle Park, N. C. 27711

April 1975

This report is issued by the Environmental Protection Agency to report technical data of interest to a limited number of readers. Copies are available free of charge to Federal employees, current contractors and grantees, and nonprofit organizations - as supplies permit - from the Air Pollution Technical Information Center, Environmental Protection Agency, Research Triangle Park, North Carolina 27711; or for a fee, from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22161.

This report was furnished to the Environmental Protection Agency by PEDCo-Environmental Specialists, Inc., Cincinnati, Ohio 45246, in fulfillment of Contract No. 68-02-1375. The contents of this report are reproduced herein as received from PEDCo-Environmental Specialists, Inc. The opinions, findings, and conclusions expressed are those of the author and not necessarily those of the Environmental Protection Agency. Mention of company or product names is not to be considered as an endorsement by the Environmental Protection Agency.

Publication No. EPA-450/3-75-065

ACKNOWLEDGEMENT

This report was prepared by PEDCo-Environmental Specialists under the direction of Mr. Charles E. Zimmer, Project Director. Mr. David W. Armentrout served as Project Manager.

We would like to express our appreciation to the following EPA offices for participating in this survey and for providing technical information:

- U.S. EPA Region I
- U.S. EPA Region II
- U.S. EPA Region III
- U.S. EPA Region IV
- U.S. EPA Region V
- U.S. EPA Region VI
- U.S. EPA Region VII
- U.S. EPA Region VIII
- U.S. EPA Region IX
- U.S. EPA Region X
- Office of Planning and Evaluation, Washington, D.C.
- Division of Stationary Source Enforcement, Compliance and Analysis Section, Washington, D.C.
- Control Systems Laboratory, Research Triangle Park, North Carolina
- Human Studies Laboratory, Research Triangle Park, North Carolina
- Strategies and Air Standards Division, Durham, North Carolina
- National Air Data Branch, Durham, North Carolina
- Office of Administration, Washington, D.C.

- Monitoring and Data Analysis Division, Durham, North Carolina
- Council on Environmental Quality, Washington, D.C.

A complete list of individuals interviewed is in Appendix C.

Mr. Gerald Nehls, National Air Data Branch, served as Project Officer for the Environmental Protection Agency. Mr. Nehls accompanied the interview team on the Regional Office visits. We wish to thank Mr. Nehls for his valuable assistance.

ABSTRACT

This report presents the results of a survey conducted among the Regional Offices of the U.S. Environmental Protection Agency and among selected Environmental Protection Agency offices in Washington, D.C., in Durham, North Carolina, and at Research Triangle Park, North Carolina. Recommendations for improvements in or additions to the current NADB systems were discussed. Specifications were written for those recommendations with sufficient available information.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENT	iii
ABSTRACT	v
LIST OF TABLES	ix
1.0 EXECUTIVE SUMMARY	1-1
1.1 Background	1-1
1.2 Approach	1-1
1.3 Survey Results	1-2
1.3.1 NEDS Recommendations	1-2
1.3.2 SAROAD Recommendations	1-3
1.3.3 Recommendations for Other Systems	1-3
1.4 Problem Areas for NEDS and SAROAD	1-3
1.5 Implementation Priorities	1-4
2.0 TECHNICAL REPORT	2-1
2.1 Background	2-1
2.1.1 Purpose and Scope	2-1
2.1.2 Interview Approach	2-2
2.2 Summary of Results	2-3
2.2.1 NEDS Recommendations	2-3
2.2.2 SAROAD Recommendations	2-14
2.2.3 SOTDAT Results	2-29
2.2.4 HATREMS Recommendations	2-29
2.2.5 SIP Rules and Regulations Recommendations	2-34
2.2.6 APER Form Recommendations	2-35
2.2.7 QAMIS Recommendations	2-36
2.2.8 Administrative Recommendations	2-37

TABLE OF CONTENTS (continued)

Page

2.3 Applications Categories 2-38

2.3.1 Data Base Operations 2-38
2.3.2 Data Analysis 2-44
2.3.3 Modeling 2-49

2.4 Implementation Priorities 2-49

2.4.1 NEDS 2-49
2.4.2 SAROAD 2-54

2.5 Implementation Requirements 2-61

2.5.1 Implementation Requirements for NEDS 2-62

2.5.2 Implementation Requirements for SAROAD 2-62

APPENDIX A TASK DESCRIPTIONS, NEDS A-1

APPENDIX B TASK DESCRIPTIONS, SAROAD B-1

APPENDIX C INTERVIEW SUMMARIES C-1

LIST OF TABLES

No.		Page
1	Survey Recommendations for Improving NEDS	2-5
2	Status of Recommended NEDS Applications In Terms of Implementation	2-15
3	Number of Users Expressing Interest in NEDS Recommendations	2-17
4	Survey Recommendations for Improving SAROAD	2-19
5	Status of Recommended SAROAD Applications In Terms of Implementation	2-30
6	Number of Users Expressing Interest In SAROAD Recommendations	2-32
7	Survey Recommendations Related to Data Base Operation	2-38
8	Survey Recommendations Related to Data Analysis	2-44
9	Benefits To Be Realized From Implementation of Applications - NEDS	2-51
10	Priority Assignments for NEDS Applications	2-55
11	Benefits To Be Realized From Implementation of Applications - SAROAD	2-56
12	Priority Assignments for SAROAD Applications	2-60
13	NEDS Data Base Operation - Implementation Requirements	2-63
14	NEDS Data Base Analysis - Implementation Requirements	2-64
15	SAROAD Data Base Operation - Implementation Requirements	2-66
16	SAROAD Data Analysis - Implementation Requirements	2-67

1.0 EXECUTIVE SUMMARY

1.1 BACKGROUND

The Environmental Protection Agency (EPA) maintains extensive computerized systems for the storage and retrieval of data on air quality measurements and air contaminant emissions. The EPA system for handling air quality data is the Storage and Retrieval of Aeromatic Data (SAROAD), and the system for handling emissions data is the National Emissions Data System (NEDS). The National Air Data Branch (NADB) is developing systems to handle other types of data such as rules and regulations data and source test data. These systems are in various stages of development. In addition, various changes to NEDS and SAROAD are planned or are currently being implemented.

The purpose of this contract was to interview EPA users of the NADB systems to determine how the current systems or planned changes and/or additions meet the users' requirements. A list of the EPA offices included in the interviews is in Appendix C. In the event that existing or proposed systems do not meet the needs of a specific user, an attempt was made to determine the specific system changes or additions that will be necessary.

1.2 APPROACH

Prior to visiting each Regional Office, a synopsis of current NADB system capabilities was forwarded to the NEDS-SAROAD coordinators for distribution among Regional Office users. Some indication of system development plans was included for each of the NADB systems, and the Regional Office users were asked to consider some pertinent questions prior to the interviews. For non-Regional Office users, an attempt was made to convey as much preparatory information as possible in telephone conversations prior to the interviews.

During each interview, the current NADB system development plans were reviewed and the survey participants were asked to comment. The survey participants were then asked to discuss specific data handling problems for which the current or projected systems are inadequate.

PEDCo reviewed all suggested changes or additions to the current systems. Where enough information exists, specifications or conceptual approaches were written, and the impact of implementation on all system elements was assessed. A schedule of priorities for implementing each system change has been suggested. The specific systems included in this contract are:

- National Emissions Data System (NEDS)
- Storage and Retrieval of Aerometric Data (SAROAD)
- Source Testing Data System (SOTDAT)
- Hazardous and Trace Elements Management Systems (HATREMS)
- SIP Rules and Regulations (SIP)
- APER Forms
- Quality Assurance Management Information System (QAMIS)
- Administrative Problems

1.3 SURVEY RESULTS

The results of this survey are summarized here in terms of system changes, additions, and administrative or operational considerations suggested by EPA users.

1.3.1 NEDS Recommendations

The major recommendations received for improving NEDS include ten (10) applications requiring new programs, three (3) applications requiring changes to existing programs, and four (4) operations changes. These recommendations are summarized in Table 1 in the Technical Report. The recommendations are related to three application categories:

- Data base operations
- Data analysis
- Modeling

A total of six (6) applications relate to improving data base operations. Eleven (11) of the applications relate to data analysis. The applications related to modeling are essentially the same as those for data analysis.

1.3.2 SAROAD Recommendations

The major recommendations received for improving SAROAD include ten (10) applications requiring new programs, nine (9) applications requiring changes to existing programs, and two (2) operations changes. Table 2 of the Technical Report summarizes these recommendations. The Technical Report relates these recommendations to the same three application categories previously mentioned for NEDS. A total of five (5) applications are related to data base operations. Seventeen (17) applications are related to data analysis. Three (3) applications relate directly to modeling capabilities.

1.3.3 Recommendations for Other Systems

No specific recommendations that can be translated into implementation specifications were received for the other systems. Most survey participants were only passively aware of these systems, since the systems either have not yet been implemented, or they have been used only marginally. Consequently, the emphasis of the survey results is on NEDS and SAROAD.

1.4 PROBLEM AREAS FOR NEDS AND SAROAD

The major problem areas for both NEDS and SAROAD related to data base operations, i.e. techniques for data input, storage, and retrieval are:

- Currency of the data
- Data quality
- Responsiveness of the system for report retrieval

Implementation of any or all of the recommendations related to each of these problem areas is expected to increase the number of data users. This in turn is expected to improve the quality of data submitted to NADB.

The major problem areas related to data analysis are:

- Data base discontinuity
- Definition of user requirements
- Quality control in data collection

The general benefit to be derived from implementing any of the recommendations associated with these three areas is an increase in the number of applications for the data. The number of users should expand as a result.

The problem areas with modeling applications are closely associated with the problems related to data analysis. Implementation of the recommendations for this application area are also expected to increase the number of users and eventually result in an improved data base.

The most important problems with operating the NADB systems are those associated with educating and communicating with the users or potential users as to system capabilities.

1.5 IMPLEMENTATION PRIORITIES

The priorities assigned for implementing each of the applications identified in the survey are based on a combination of four general expected benefits:

- Improvement of data quality
- Reduction of input/output time and report turnaround time
- Savings of man-hours
- Improved usage of the systems

Table 10 of the Technical Report summarizes the priority assignments for the NEDS application; Table 12 of the Technical Report summarizes them for SAROAD applications. For both NEDS and SAROAD, the highest priorities are on the applications that improve procedures for updating the data bases and for retrieving reports.

2.0 TECHNICAL REPORT

2.1 BACKGROUND

2.1.1 Purpose and Scope

This survey was conducted among all EPA Regional Offices and thirteen (13) selected EPA offices at the Research Triangle Park, North Carolina and in Washington, D.C. The offices visited and the names of the coordinating contacts in each office are listed in Appendix C. The purpose of the survey was to determine if the National Air Data Branch (NADB) systems are meeting the current and projected needs of the users within EPA.

Where-in it was determined that existing systems can be changed to be more responsive to EPA users' data requirements, specifications for such changes are defined. In some instances, not enough information was available to evaluate the benefit derived by the implementation of an application. In these cases more information is needed from the potential users to justify implementation. For other cases, the benefits to be derived are obvious and they represent solutions to problems that were discussed in the report "Establishment of a Non-EPA User System for State Implementation Plans" (Contract No. 68-02-1001, Task 4).

The specific topics discussed in the survey were:

- National Emissions Data System (NEDS)
- Storage and Retrieval of Aerometric Data (SAROAD)
- Source Testing Data System (SOTDAT)
- Hazardous and Trace Elements Management System (HATREMS)
- SIP Rules and Regulations (SIP)
- APER Forms
- Quality Assurance Management Information System (QAMIS)
- Administrative Problems

The survey recommendations under each topic are categorized as additions to existing systems, changes to existing systems, or operations changes. Where sufficient information exists, applications have been detailed in the Appendices. Each recommendation is analyzed in terms of implementation cost, and impact on the total system.

General benefits that could be derived from the applications presented in this report are:

- ° Increase user participation and interest in the systems.
- ° Improve the quality of data and the currency of the data in the data bases.

These benefits complement each other, i.e. as user participation increases, the quality of data can be expected to improve and vice versa. For any of the recommendations to be effective, significant effort for improved communications between NADB, the Regional Offices, and the states is essential. The availability of any one of the capabilities identified here will not in itself improve the system. All users or potential users will need to be educated as to the benefits to be derived from using the system. Until this is accomplished the overall quality of the data bases may not show any significant improvement. The education and communication problems cannot be overemphasized - they are the most serious problems associated with operating the NADB systems. These problems are discussed in detail in Section 2.2 in this report.

2.1.2 Interview Approach

For each office included in the survey, a representative was asked to coordinate with other users in the office to include them in the interviews. This was an attempt to obtain information concerning as many systems as were being used within each office and to incorporate the views of users with as many different applications as possible.

Prior to each interview in the Regional Offices, a synopsis of current NADB system development plans, and a summary of the problems to be addressed in the interview was mailed to the Regional Office AEROS contacts.

A sample of the pre-interview materials is included in Appendix C. During each interview, the tentative system development plans prepared by NADB were presented, and comments on each application were solicited. Each user was then asked to discuss their special data requirements. An attempt was made to relate those needs to current system capabilities. When a user indicated that current and planned systems did not fulfill specific data requirements, detailed specifications for such requirements were requested.

In many cases users were not familiar with all of the options of the various systems. In most cases those interviewed were not familiar enough with the systems to be specific in discussing their data needs in relation to the quality of the data bases and the responsiveness of the systems. The results of each Regional Office interview were summarized and sent to the AEROS contacts for review and comment. The Project Officer was also provided with a copy of the results of each interview, and he has contacted each AEROS coordinator to discuss any applications or problems not covered in this report.

2.2 SUMMARY OF RESULTS

This section is a synopsis of the user recommendations for improving each of the NADB systems discussed in the survey. For many recommendations, not enough information was provided to directly justify implementation. In these cases, it is incumbent on the users, especially the Regional Offices, to discuss the recommendations and to provide NADB with further justification for implementing the applications of interest.

2.2.1 NEDS Recommendations

The survey recommendations for improving the NEDS include ten (10) applications requiring new programs, three (3) applications requiring changes to existing programs, and four (4) operations changes. The applications are summarized in Table 1 and cross-referenced to Appendix A for detailed discussion where necessary.

2.2.1.1 NEDS New Program Applications -

The applications discussed here require new programs or subroutines. In some cases, the applications necessitate changes in several other programs or files. The total impact of each application on the other system components is indicated in the applicable sections of Appendix A.

2.2.1.1.1 Latitude - longitude input with subsequent UTM conversion - This application would allow NEDS users to input latitude-longitude coordinates instead of UTM coordinates. The recommendation is based on the fact that many states routinely use latitude-longitude coordinates, and they consequently view the job of providing UTM coordinates to NEDS as unnecessary. As a result, the locator data for many state emission inventories is less than desirable. The result is reflected in increased time and manpower required for any updating or modeling tasks. The capability for inputting latitude-longitude coordinates directly into NEDS should be a separate procedure from the normal activities involved with completing NEDS coding sheets. Inclusion of a separate field on the coding sheets requires a change of the NEDS form, changes to several programs, restructuring of the master record, and rebuilding of several files. Using a modified version of the TCLCONV program that converts latitude-longitude to UTM's and then submitting the UTM coordinates as updates to the NADB* NEDS-USER file has several advantages. The TCLCONV program can be changed easily to produce punched card output in NEDS format. The update of the USER file will occur as part of the normal update activities. The only impact on NADB's operations will be the requirement for a programmer to make the minimal change to TCLCONV.

2.2.1.1.2. Polygon - defined area retrieval- This application would allow a user to define, by latitude/longitude coor-

Table 1. SURVEY RECOMMENDATIONS FOR IMPROVING NEDS [a]

New programs	Changes to existing programs	Operations changes
1. Latitude-longitude input with UTM conversion (A1)[b]	11. Expanded comments capability	14. Implement user training seminars (A8)
2. Polygon retrieval (A2)	12. Indicate sources that have been deleted from a report because of confidentiality (A7)	15. Discontinue or modify the Verification Report
3. Retrieve by range of parameter (A3)	13. AQMA retrieval	16. Add fugitive dust SCC's (A9)
4. Analyze the effect of potential changes to a parameter on NEDS data (A4)		17. Improve the format for identifying report retrieval options (A10)
5. Lowspeed terminal update		
6. Usage statistics by report and/or user		
7. Basic accounting capability (A5)		
8. Emission trends		
9. Decentralization of edit		
10. NEDS-CDS cross-reference (A6)		

[a] Additional applications i.e., those mentioned in only one interview, are included in Appendix C.

[b] Refers to the Appendix section where application specifications are found.

dinates, a polygon shaped geographical area and to retrieve information on all sources meeting specified criteria and lying within the polygon area. The user would also have the capability to retrieve a list of NEDS points or SAROAD active sites within a specified distance from a point in the polygon. The capability to retrieve a listing of active SAROAD sites within a polygon area would also be included. Implementation of this application requires a feasibility study because current NEDS retrieval options are flexible enough that, if used properly, they might produce the same results. This application might be useful for AQMA retrieval.

2.2.1.1.3 Define a value range for a parameter and retrieve sources with parameters in the range - This application allows users to specify a value range for any data element in the NEDS record and retrieve only those sources within the defined range. An example is the retrieval of all coal boilers (controlled by SCC) within a county (controlled by County Code) that burn coal between 2.5 and 3.5 percent sulfur The expected benefits from this application are similar to those for polygon - defined area retrieval (Section 2.2.1.1.2). This application is more directly useful for special studies than for general application, and users such as Control Systems Laboratory can provide the major input for a feasibility study. This application is particularly well suited to file management systems such as System 2000. A secondary benefit from this application might be to save users from writing special application programs for batch processing using NADB files. Moreover, system turnaround time will be improved if this application is available in an interactive mode.

2.2.1.1.4. Analyze the effect of potential changes to a parameter - This application allows users to substitute new values for parameters in an existing point source record and to analyze the resulting effects on pollutant emissions. The major requirement for this capability is for control strategy testing and Air Quality Maintenance Area plans development. For example, the ability to change sulfur contents allows users to evaluate the effect of requiring all sources of a specific type to burn fuel with percent sulfur less than some specified maximum value. This application is easily handled by a file management system and will save a significant amount of time in preparing data for modeling.

2.2.1.1.5. Lowspeed terminal update capability - This application enables NEDS users to interactively edit new data. Because of the cost associated with interactive operation it is appropriate for less than 50 punched cards. Upon passing edit, the data are added to a temporary file that is accessed by NADB to update the NADB*NEDS - USER file. This capability will allow Regional Offices to input data rejected by a previous edit run into the NEDS system immediately after correction. Such an update capability will reduce the effort on the part of NADB, provide more direct control over updates by the Regional Offices, and theoretically, result in a more current data base than is presently available. Success of this application requires the following: 1) NADB must establish a minimal update schedule to assure that data entered to the update file are added to the USER file in a timely manner; 2) an audit system must be developed to notify the Regional Offices that a maximum specifed time period has passed between the last edit and the reentry of any cards rejected in the last edit.

2.2.1.1.6. System usage statistics by report type - This application provides users with a monthly or quarterly report of the usage of specific NEDS report generating programs. This can be a valuable budgeting tool for users. It may also indicate application areas which are used infrequently and might require changes to make them more applicable to user needs. It can be a valuable planning tool for NADB to determine areas or applications for which user activity is increasing or decreasing. For this particular survey, for example, historical information on usage by report type would have been valuable for helping to establish implementation priorities. This application requires a feasibility study. It should be directed to the EPA computer center at Research Triangle Park, since generation of the required information is their responsibility.

2.2.1.1.7. Basic accounting capability - This application allows a user to determine the number of sources or facilities in the file that conform to some defined criteria, e.g. the number of coal-fired boilers in the file or the number of sources with boilers burning multiple fuels. The users can then determine whether or not there is sufficient data in the file to warrant a particular report retrieval. The expected result is that manpower is saved in scanning reports and system time is saved by not retrieving information which serves no useful purpose for the user. This application requires a feasibility study, since the expected benefits are based on speculation on the part of the potential users. This application lends itself to an interactive mode file management system.

2.2.1.1.8. Emissions trends by area and by source type - This application provides an analysis of data over a period of years to show general trends in emissions for a specific pollutant, and it is especially useful for AQMA plan development. Immediate implementation of this capability is questionable for two reasons:

- Much of the data in NEDS are not updated regularly. Consequently, significant gaps in the data for each year exist. A schedule for updating operating parameters and emissions estimates in NEDS needs to be enforced. Implementation of automated emissions inventory systems among the states and/or increased usage of NEDS by the states would be expected to result in increased cooperation by the states in submitting their semiannual updates.

- Several states have elected to replace their entire inventories in NEDS with revised inventories. Consequently, a direct comparison of inventories within an affected state is impossible.

2.2.1.1.9. Decentralization of the edit programs - This application shifts the responsibility for initial edit of data from NADB to the Regional Offices. A feasibility study for this application has been completed and NADB is proceeding with implementation. Consequently, no indepth analysis is attempted in this report. The major benefit expected from this effort is to decrease the time required for updating the system. The current update mechanism requires the Regional Offices to submit data to NADB for edit. NADB then sends the edit results to the Regional Offices, where the data is either corrected or sent to the appropriate state for correction. The data are then resubmitted to NADB for another edit, validation and file update. It is anticipated that from two to four weeks can be cut from the current update time. This capability, coupled with the capability for lowspeed terminal data entry the Regional Offices will significantly improve the currency of the data base. The update time now required for NEDS has been offered by several states as a major reason for not using NEDS.

2.2.1.1.10 Include a CDS cross-reference number in NEDS - This application allows NEDS users, primarily Regional Offices to include in each point source record an identifier to reference that point in CDS. This capability will allow NEDS and CDS users to determine the degree to which data in each data base can augment the other. The cross-reference number from CDS could be entered into a separate comments card for each point source (see Section 2.2.1.2.1). A separate cross-reference file is required to indicate updates that have occurred in CDS for which input data to NEDS might be required. A program to update the cross-reference file and to produce cross-reference listings is required. Implementation of this application requires a committment on the part of the Division of Stationary Source Enforcement (DSSE), NADB, and the Regional Offices to perform the initial cross-reference against CDS and to keep both systems current to reflect updates. This application warrants a feasibility study.

2.2.1.2 NEDS Program Changes -

These applications require changes to existing programs. The total impact of each change on other system components is indicated in the respective section of Appendix A when necessary.

2.2.1.2.1. Expanded comments fields - This application is being implemented by NADB. The capability allows users to input one or more Card 7 for each plant in the file. The Card 7 contains the plant, point, year, and SCC (if applicable) identifiers to parallel the other six cards. The remainder of the Card 7 is used to enter any pertinent comments. One Card 7 is allowed for general plant information; one Card 7 is allowed for each point source; and one Card 7 is allowed for each SCC within a point source. Implementation of this capability resulted partially from the need for users, especially DSSE users, to see equipment identification data associated with each NEDS record. Also, this capability fulfills the need for providing cross-reference information for state permit systems or for state emission inventory systems. This can provide NEDS users with the ability to trace data in NEDS back to original documents. Acquisition of the necessary information will require significant cooperation from the states.

2.2.1.2.2. Indicate confidential sources on reports - This application will allow users the option of printing plant and point identification numbers and name and address for sources that have not been included on a report because of confidentiality. Currently, if a point source is coded as confidential, the record for the point source is skipped during program execution. Consequently, the results in the report do not reflect the contributions from the confidential sources. No indication is given that a source has been excluded because of confidentiality. Implementation of this capability is relatively inexpensive and no significant impact on the total system is expected. Indications are that NADB will await the Office of Enforcement and General Council (OEGC) ruling on confidentiality before formalizing procedures for this.

2.2.1.2.3. AQMA report retrieval - The purpose of this application is to allow NEDS users to request reports by AQMA number. The request mechanism works in the same fashion as the current mechanisms for state, county, and/or AQCR retrievals. AQMA retrieval will have the following impact on the NEDS system:

- The NEDS coding form must be restructured, requiring OMB clearance.
- AQMA codes must be published.
- A minimum of four files must be reformatted.
- A minimum of fifty programs must be recompiled.
- AQMA codes must be added to approximately 100,000 facilities currently in the NEDS system.

The original intent of the guidelines for defining AQMA's was that they should follow political jurisdiction boundaries. If AQMA's are defined in this manner any selected report can be produced simply by specifying the appropriate counties. If AQMA boundaries do not follow political jurisdictions, the polygon - defined area retrieval capability could satisfy the requirement for AQMA retrieval. Because of the high cost of implementing AQMA retrieval, and the fact that existing options can satisfy most retrieval requests this application is not justified at this time. An alternative approach could be to set up a cross-reference file of AQMA numbers versus SAROAD county numbers, or AQMA numbers versus UTM coordinates. This approach should be investigated before any decisions are made concerning AQMA retrieval.

2.2.1.3 NEDS Operations Changes -

This section includes user recommendations for changes to NEDS operating procedures.

2.2.1.3.1. Implement user training seminars - NADB has conducted several NEDS user seminars. Reactions from the Regional Offices regarding the effectiveness of the seminars is favorable, and several Regional Offices have requested that they be offered routinely. The seminars offered thus far have covered these topics:

- NEDS point source coding, updating, and edit/validation procedures.
- NEDS area source data development and coding.
- Summary of NEDS output formats.
- Introduction of the Emissions Inventory Subsystem (EIS).
- EIS coding and update procedures. A training manual has been prepared to supplement existing training materials. The seminars appear to be effective for solving problems regarding data input. If they are continued, they can be expected to have a long-term positive effect on the quality of data being input to NEDS.

An additional area of instruction that should be presented in seminars is the topic of how to use the NEDS system. Most states seem to be unaware of the system capabilities. Moreover, they have been frustrated by the time delays in getting data into and out of the system. The states are primarily familiar with NEDS because of the requirement for them to submit semiannual updates. If users, especially state agencies, are made aware of benefits to be derived from using the NEDS system, and if they can be made aware of all of the options for data retrieval and analysis, a significant increase in the use of the system should occur. Increased use of the system would be expected to generate more interest in the quality and currency of the data going into the system. The result should be a more reliable data base.

User seminars should emphasize both management and engineering applications for NEDS data. A survey of EPA users of NEDS should be conducted to determine their specific applications prior to determining the content of the seminars. Results of the survey should be reviewed to determine how the EPA applications might be related to state agency operations. These recommendations are supported by the previous user survey conducted among selected state agencies as well as by the current survey.

2.2.1.3.2. Modify the verification report and validation listings - Five Regional Offices indicated that handling and reviewing the Point Source Listing produced from the Verification file is cumbersome. Often the listing is ignored.

The Validation Listing produced from the NEDS edit routines was also mentioned as a topic of concern in the survey. Current efforts for decentralizing the edit/validation routines to the Regional Offices include plans to produce point source listings only for those plants with parameters exceeding the allowed validation values. The point source listing is essential to validation procedures in order to save effort required from the states. A point source listing further enables the Regional Office to discriminate on specific sources for which the state should supply information. Consequently, the listing should continue to be produced.

2.2.1.3.3 Add fugitive dust SCC's - This application will result in the addition of new SCC codes to the system. The new SCC's will represent selected fugitive dust sources such as coke piles for steel plants. The recommendation of any new SCC of this type should be a Regional Office responsibility, since many of the industry types involved are regionally oriented. Because of the cost involved with developing an emission factor for each new SCC, preliminary studies should be conducted to address the following problems.

- Current emission factors for the industry of interest must be investigated to insure that the specific fugitive dust sources of interest have not already been included.

- A decision must be made as to whether the fugitive dust category could be considered a point source, or if it would better qualify as an area source.

- An estimate should be made of the probable impact of the fugitive dust source in relation to the total particulate emissions for several plants or processes representative of the industry.

2.2.1.3.4 Improved format for defining retrieval options - NADB has printed explanations of keying options for report retrieval, and they have been widely circulated. Many users have indicated that these explanations are cumbersome and difficult to be understood by anyone other than systems - oriented personnel. It has been suggested that a simplified reference table format would be more easily interpreted. An example format is included in Appendix A. The format lists reports by name and indicates by 'yes' or 'no' if each report is available by specific retrieval options. Circulation of a similarly formatted matrix table, preferably with sample reports attached, will make potential users aware of benefits to be derived from accessing the system. The AEROS contacts constitute the largest group of users who understand the retrieval options.

2.2.1.4. NEDS Recommendation Summary -

The previous discussions for NEDS recommendations resulting from this survey are summarized in Table 2 in terms of requiring feasibility study or being ready for immediate work toward implementation. The assessment is based solely on the amount of background information available at the time of the survey. Applications currently being implemented are so indicated. Table 3 summarizes the recommendations in terms of the number of users who expressed interest.

2.2.2 SAROAD Recommendations

The survey recommendations for improving the SAROAD system include ten (10) applications requiring new programs, nine (9) applications requiring changes to existing programs, and two (2) operations changes. The applications are summarized in Table 4 and cross-referenced to Appendix B for detailed discussion where applicable.

2.2.2.1 SAROAD New Program Applications -

The applications discussed here require new programs or subroutines. Some applications necessitate changes in several other programs or files. The total impact of each application

Table 2. STATUS OF RECOMMENDED NEDS APPLICATIONS IN TERMS OF IMPLEMENTATION

No.	Application	Requires feasibility study	Ready for implementation planning	Implementation under way
1.	Latitude-longitude with UTM conversion	X		
2.	Polygon retrieval	X		
3.	Retrieval by range of parameter	X		
4.	Analyze effect of potential changes to parameters in source record	X		
5.	Lowspeed terminal update		X	X
6.	Usage statistics by report and/or user	X		
7.	Basic accounting capability	X		
8.	Emission trends	X		
9.	Decentralization of edit	(completed)		X
10.	NEDS/CDS cross-reference	X		
11.	Expanded comments capability		X	X
12.	Indicate sources deleted from report because of confidentiality		X	
13.	AQMA retrieval	X		
14.	Implement user training seminars		X	(partial)

Table 2 (continued. STATUS OF RECOMMENDED NEDS APPLICATIONS IN TERMS OF IMPLEMENTATION

No.	Application	Requires feasibility study	Ready for implementation planning	Implementation under way
15.	Discontinue or modify Verification Report	X		
16.	Add fugitive dust SCC's	X		
17.	Improve format for identifying report retrieval options		X	

Table 3. NUMBER OF USERS EXPRESSING INTEREST IN NEDS RECOMMENDATIONS

No.	Application	Interested users			Total
		Regional offices	RTP	Washington	
1.	Latitude-longitude with UTM conversion	3	1		4
2.	Polygon retrieval	2	1		3
3.	Retrieval by range of parameter	5	2		7
4.	Analyze effect of potential changes to parameters in source record	2			2
5.	Lowspeed terminal update	8		1	9
6.	Usage statistics by report and/or user	1			1
7.	Basic accounting capability	2	1		3
8.	Emission trends			2	2
9.	Decentralization of edit	8		1	9
10.	NEDS/CDS cross-reference	5		1	6
11.	Expanded comments capability	5	1		6
12.	Indicate sources deleted from report because of confidentiality	3			3
13.	AQMA retrieval	10	2		12
14.	Implement user training seminars	3			3

Table 3 (Continued). NUMBER OF USERS EXPRESSING INTEREST IN NEDS RECOMMENDATIONS

No.	Application	Interested users			Total
		Regional offices	RTP	Washington	
15.	Discontinue or modify verification report	5			5
16.	Add fugitive dust SCC's	2			2
17.	Improve format for identifying report retrieval options	2	1		3

Table 4. SURVEY RECOMMENDATIONS FOR IMPROVING SAROAD[a]

New programs	Changes to existing programs	Operations changes
1. Polygon retrieval	11. Option to include data not meeting 75% criteria (B5)	20. English language retrieval
2. Trends plotting (B1)[b]	12. Include 2nd maximum on Standards Report (B6)	21. Standards Report more frequently
3. Audit SIP station reporting (B2)	13. Identify inactive sites (B7)	
4. Lowspeed terminal update	14. Include reporting units on edit and validation reports	
5. Usage statistics by report and/or user	15. County retrieval (B8)	
6. Graphics to plot site locations (B3)	16. Include minimum detectable levels on reports (B9)	
7. Calculate wind rose and pollutant rose	17. Include accuracy of method in Site I.D. File (B10)	
8. Parametric data retrieval	18. AQMA retrieval	
9. Report on Standards violations (B4)	19. Reduce conversation with interactive retrieval	
10. Decentralized edit		

[a] Additional applications i.e., those mentioned in only one interview, are included in Appendix C.

[b] Refers to Appendix section where application specifications are found.

on the other system components is indicated in each respective section of Appendix B.

2.2.2.1.1. Polygon - defined area report retrieval - This application is part of the general purpose retrieval package previously discussed for NEDS (Section 2.2.1.1.2).

2.2.2.1.2. Trends plotting - The purpose of this application is to allow graphic presentation of the trends in air quality over a specified interval of time at a monitoring site. All plotted data could be shown in relation to the annual standard as the baseline. The 12 month moving arithmetic or geometric means for the pollutants of interest are also desirable. One recommendation called for using the projected annual maximum concentrations as the baseline with the emergency episode level as an upper limit. Many Regional Offices are currently producing trends graphs manually. A significant savings in manpower and more widespread dissemination of trends data could be expected to result from implementation of this capability.

A feasibility study is required for this application in order to determine the specific plotting requirements that would be of widespread interest. A significant problem associated with this application is that agency/project codes have changed for many monitoring stations with no concurrent traceability throughout the SAROAD system. The problem (See Section 2.2.2.2.3) should be resolved prior to implementation of the trends analysis capability. Serious misinterpretation of data could result otherwise. A second problem is the continuity of quality control procedures for each site over a long time period (2.2.2.2.7). Major discrepencies in the quality control procedures could leave doubt as to the validity of any trends analyses.

2.2.2.1.3. Capability to track SIP station reports - The purpose of this application is to provide the Regional Offices with the capability to track data reported by the states for SIP required sites. At least six Regional Offices use a manual logging system to keep track of this information.

The estimated time expended is 2 1/2 man-days/quarter/state. A significant savings in time could be realized from implementation of this capability. Moreover, implementation is straightforward, and most elements in SAROAD would not be impacted.

2.2.2.1.4. Lowspeed terminal update - This application will enable SAROAD users to interactively edit small quantities (probably fifty or fewer) of input cards. The definition and technical considerations for this application are the same as for the similar application for NEDS (See Section 2.2.1.1.5).

2.2.2.1.5. Usage statistics by report type - This application is the same as for NEDS (See Section 2.2.1.1.6). Implementation of this application will require a feasibility study directed to the EPA Computer Center at Research Triangle Park.

2.2.2.1.6. Graphics to plot site locations - This application allows SAROAD users to access the Site File and generate a plot of site locations within a designated area. The options for implementing this capability are to produce a relative location type plot on-line on the printer or to produce a Cal-Comp plot off-line. An on-line plot is relatively inexpensive and quick. Off-line plots have greater flexibility in that they could include an outline of the area of interest. The cost of off-line plots is slightly higher and the turnaround is not as rapid as for on-line printer plots. A feasibility study is required to determine the benefits to be gained by showing the area plot. An additional consideration for this application is that it may be directly related to any trends plotting capabilities as a first step in the process.

2.2.2.1.7. Calculate wind rose and pollutant rose - This capability provides SAROAD users with reports on co-analysis of air quality data and meteorological data collected at specific sites. Users will be able to more accurately assess the applicability of individual sites for modeling purposes, since an enhanced capability to resolve anomalous air quality data will be gained. Implementation of this capability will provide a combination of printed and plotted outputs. This capability does not duplicate, but rather it enhances reporting efforts available through the National Oceanic and Atmospheric Administration (NOAA). A feasibility study on this application has been completed.

2.2.2.1.8. Parametric data retrieval - This application allows SAROAD users to select and retrieve data from a variety of SAROAD files. The selection criteria are based on approximately 40 parameters or on any combination of those parameters. The usefulness of parametric data retrieval has been well established through prior use of MARK IV on the IBM system. The problem at this point is concerned with deciding on the most desireable of two approaches: 1) implementation of a file management system with the same capabilities as MARK IV or; 2) development of a COBOL program to perform the equivalent functions. Problems exist with the ability of System 2000 to handle multiple files. This problem has led to the investigation of other file management systems. Investigation of the specific MARK IV applications used by Regional Offices in addition to those used at the NERC, Research Triangle Park should be initiated before beginning any development of COBOL programs. This is necessary to insure that the program or programs involved provide sufficient flexibility so that no net loss from the capabilities on the previous IBM system is incurred.

2.2.2.1.9. Report on standards violations - This application produces a report similar in format to the Yearly Summary Report. The report will reflect only those statistics necessary to compare air quality standards with pollutant levels measured at a given site. It will be available for the criterion pollutants. This application is recommended highly by Regional Offices. Related recommendations concerning percent of time the standards are violated have been made by the Monitoring and Reports Branch. Implementation of this capability will not significantly impact the SAROAD system. A feasibility study is not necessary, since prior communications adequately describe the need for this application.

2.2.2.1.10. Decentralized edit - This application parallels the same application mentioned for NEDS. Implementation is in progress at NADB.

2.2.2.2 SAROAD Program Changes -

The applications discussed here require changes to existing programs. The total impact of each change on other system components is indicated in the respective section of Appendix B.

2.2.2.2.1. Option to include data not meeting the 75% criteria in Summary Reports - Currently, 75% of the possible data values from a particular station for a summary period must be available before the summary statistics are calculated and stored on the appropriate summary file. The capability to calculate summary statistics regardless of the amount of data available can be implemented by changing the existing program logic. If this is done, summary statistics based upon data not meeting the 75% criteria will be so flagged and reported or included in subsequent statistical calculations only when so requested by the user. This approach protects the interest of any users desiring to see only statistics for data meeting the 75% criterion. Routine reports should be based upon the 75% completeness in order to maintain a minimal level of quality control without requiring major reorientation of the users. This approach requires users desiring this option to be aware of the deficiencies in the data used to calculate the summaries. The Regional Offices should be asked to respond on the advantages of this option.

2.2.2.2.2. Include 2nd maximum values on the Standards Report - This application provides users with the second highest value for the sampling or averaging period of interest. Currently, the system prints the maximum observed values on the Inventory Report, on the Yearly Report by Quarters and on the Quarterly Frequency Distribution. The second maximum observed value is necessary to determine if the air quality standard has been violated. As sampling frequencies increase so will the effort required to manually find the second maximum value. Implementation of this capability is justified, because of the relatively low cost of implementation and the potential for saving manhours.

2.2.2.2.3. Identify inactive sites - This application requires that an active/inactive code be added to each record in the Site File so that only data for active sites can be retrieved. A net saving of computer time and manpower for reviewing reports will result. The problem, however, goes beyond merely identifying each site as active or inactive. When there is a transfer of agency responsibilities related to site operation, the Agency Code changes. The effect on the system is that a new record is created since the Agency Code is a key identifier for each site. Consequently, new data are stored in the system according to the new Site Code. The data under the old Site Code are not applicable for summary statistics, etc. performed on data submitted under the new code. Although the same station is physically active, the system considers the station as defined under the previous code as being inactive. The loss in data continuity causes problems for users interested in trends analyses and/or in tracking the data submittal status for a site. One major effort has already been expended in which contractors nationwide performed a manual search for missing data in the SAROAD files and attempted to clarify the status of sites in the SAROAD system.

Unless NADB provides a method to allow tracking changes of site code within the system the same effort may be required again. The problem should be solved as soon as possible. A suggested approach from the Monitoring and Reports Branch (MRB) of the Monitoring and Data Analysis Division includes the use of four status codes: 1) active site; 2) inactive site; 3) modified active site; 4) modified inactive site. These codes in conjunction with initial and terminal dates could provide the necessary tracking capability to provide the following advantages:

- Knowledge of all sites and their status
- Ease of relating changes in agency or projects to the effect on air quality data
- More summary statistics, since changes during the year would be reflected in annual computations
- Retrieval and data manipulation relating to a specific location would be easier

These comments from MRB provide the most indepth approach to this problem offered during the survey. The majority of the Regional Offices expressed concern with this problem. Although a significant amount of processing would be required to rework the files, and an initial large effort might be required to input the necessary historical site information, the long-term benefits for data analysis and cost effectiveness seem to warrant implementation of this application.

2.2.2.2.4. Show reporting units on edit and validation reports - This recommendation will save time in reviewing edit and validation reports. State agencies who want to check results currently must convert the raw data values to reporting units first. The reporting units could be shown on the reports with minimal programming changes.

2.2.2.2.5. County retrieval - This application allows SAROAD users to retrieve reports for sites within county. This capability will reduce the time necessary to review reports in addition to saving machine processing time. County retrieval can fulfill user requirements for AQMA retrieval if AQMA's follow county boundaries. Implementation requires relatively little effort, since the county code is already in the Site File.

2.2.2.2.6. Include minimum detectable levels on reports - This application allows users to compare the observed values on any report to the minimum detectable level for the sampling method. This comparison could provide a better idea of overall data quality by making averaging biases more visible to users. The minimum detectable level will be included in the report heading. The system impact for this application is relatively minor and further investigation should not be necessary.

2.2.2.2.7. Include statement on method accuracy in the site identification file - This application will allow users to see a statement of the method accuracy on each report. The users can then better assess the data reliability as related to a specific project need. The application is relatively easy to implement, but it requires an expansion of the Site Record.

The recommendation for this application points out the need for a much more involved capability to qualify SAROAD data in terms of quality assurance information. Depending on the volume and type of quality assurance information needed, and on the requirements for reporting this information, it may be necessary to significantly revise the SAROAD system. The information necessary to evaluate data quality should be defined as soon as possible.

2.2.2.2.8. AQMA retrieval - This option which allows SAROAD users to request reports by AQMA number has the following impact on the system:

- The SAROAD coding form must be restructured, requiring OMB clearance.
- AQMA codes must be published.
- Most SAROAD files must be reformatted,

- Most SAROAD programs would require recompilation.
- AQMA codes must be added to all site identification codes.

If AQMA's are assigned according to political jurisdiction boundaries, as was the original intent of the AQMA guidelines, then county retrieval or polygon area retrieval could complement current retrieval options sufficiently for AQMA applications. Because of the uncertainity of how AQMA's will be designated, and because of the high cost of implementation, this application is not justified at this time.

2.2.2.2.9. Reduce the conversational aspects of interactive Retrieval - The current procedure for conversational retrieval requires the user to answer questions at key points such as at the end of a record on the Site Description Report. At the end of each site description, for example, the user is required to enter the number for the next site. The user essentially is tied to the console until the end of job in this case. The recommendation for reducing the conversational aspects allows terminal users to enter most of the retrieval keys and selection criteria only one time at the beginning of a run. Because the user will not be required to constantly enter replys to the system, his attention can be diverted to other tasks while the output to his program is being printed. Implementation of this capability should be an option, not a user requirement, since for some retrieval requests, the user's desire for more data may depend on the information printed in the last data block.

2.2.2.3. SAROAD Operations Changes - This section includes user recommendations for changes to SAROAD operating procedures.

2.2.2.3.1. Write a user manual to allow English language retrieval - This recommendation applies primarily to users outside of the Regional Offices. The recommendation reflects the fact that SAROAD users are now required to reference a series of manuals to obtain the necessary information to enter all codes necessary to retrieve a report. Understanding the codes in the SAROAD retrieval key is necessary to properly use the system. Because of the large number of code combinations available, English language retrieval still requires a user to reference parameter tables, and the chance of entering the wrong combination can be expected to be as great as they are with numeric code retrieval. Implementation of English language retrieval will require revision and redistribution of at least three user's manuals. Significant programming effort will be required, and table lookups required for each run will require more computer time. A more reasonable approach to English retrieval might be to revise the "Terminal User's Manual" to include English language examples showing the origin of the resulting retrieval codes. This approach can be partially documented and circulated to users for comment before implementation. Any consideration of English language retrieval should be preceeded by a feasibility study.

2.2.2.3.2. Generate standards report more frequently- This application allows users, especially Regional Office users, to monitor compliance with air quality standards for a specified area on a more timely basis. Currently, the Standards Report is generated quarterly. Current NADB plans include generating the Standards Report each time the SAROAD data are updated.

2.2.2.4 Recommendation Summary -

The previous discussion for SAROAD recommendations resulting from this survey are summarized in Table 5 in terms of requiring feasibility study or being ready for immediate work toward implementation. The assessment is based solely on the amount of background information that was available at the time of the survey. Applications currently being implemented are also indicated. Table 6 summarizes the recommendations in terms of the number of users who expressed interest.

2.2.3 SOTDAT Results

The SOTDAT system was discussed briefly with each Regional Office. All comments on this system were speculative, because none of the survey participants had received information other than what had been initially circulated by NADB in a brief brochure. A follow-up correspondence survey might be warranted for this system after example reports have been circulated to the Regional Offices and to the states. The comments received for this system were:

- A status report indicating the number and types of tests in the system should be circulated periodically. Circulation to the states as well as to the Regional Office might generate state interest and consequently increase participation.
- The office responsible for each test should be noted on each report so that users could contact that office for more information.
- Data should be available by source category.

The Control Systems Laboratory was the only office within the NERC, Research Triangle Park, that indicated possible applications for this system. No specific comments were made.

2.2.4. HATREMS Recommendations

No significant comments were made regarding HATREMS. The majority of the Regional Offices indicated that the system is of no advantage to them because of the insignificant number of hazardous sources or because they had no firm needs for inventories of other pollutants. Both the Human Studies Laboratory and Control Systems Laboratory indicated that they might have applications for the system but they were unable to make specific comments.

2.2.5. SIP Rules and Regulations Recommendations

The SIP Rules and Regulations System was discussed with the Regional Offices and with the Control Systems Laboratory. The Control Systems Laboratory representative suggested that

Table 5. STATUS OF RECOMMENDED SAROAD APPLICATIONS IN TERMS OF IMPLEMENTATION

No.	Application	Requires feasibility study	Ready for implementation planning	Implementation under way
1.	Polygon retrieval	X		
2.	Trends plotting	X		
3.	Audit SIP station reporting		X	
4.	Lowspeed terminal update			(planned)
5.	Usage statistics by report and/or user	X		
6.	Graphics to plot site locations	X		
7.	Calculate wind rose/ pollutant rose	(completed)		X
8.	Parametric data retrieval		(problems to be resolved)	
9.	Report on standards violations		X	
10.	Decentralized edit			X
11.	Option to include data not meeting 75% criteria		X	
12.	Include 2nd maximum on Standards Report		X	
13.	Identify inactive sites		X	
14.	Include reporting units on edit and validation reports		X	
15.	County retrieval		X	
16.	Include minimum detectable levels on reports		X	

Table 5 (Continued). STATUS OF RECOMMENDED SAROAD APPLICATIONS IN TERMS OF IMPLEMENTATION

No.	Application	Requires feasibility study	Ready for implementation planning	Implementation under way
17.	Include accuracy of method in site I.D. file		X	
18.	AQMA retrieval	X		
19.	Reduce conversation with interactive retrieval	X		
20.	English language retrieval	X		
21.	Standards Report more frequently		X	X

Table 6. NUMBER OF USERS EXPRESSING INTEREST IN SAROAD RECOMMENDATIONS

No.	Application	Interested users			Total
		Regional offices	RTP	Washington	
1.	Polygon retrieval	3	1		4
2.	Trends plotting	7	1		8
3.	Audit SIP station reporting	6	2		8
4.	Lowspeed terminal update	1	1	1	3
5.	Usage statistics by report and/or user	1			1
6.	Graphics to plot site locations	2		1	3
7.	Calculate wind rose/pollutant rose	3	1		4
8.	Parametric data retrieval		1		1
9.	Report on standards violations	1		1	2
10.	Decentralized edit	10			10
11.	Option to include data not meeting 75% criteria	4	2		6
12.	Include 2nd maximum on Standards Report	3	2		5
13.	Identify inactive sites	6	2		8

Table 6 (continued). NUMBER OF USERS EXPRESSING INTEREST IN SAROAD RECOMMENDATIONS

No.	Application	Interested users			Total
		Regional offices	RTP	Washington	
14.	Include reporting units on edit and validation reports	3			3
15.	County retrieval	7	1		8
16.	Include minimum detectable levels on reports		1		1
17.	Include accuracy of method in site I.D. file		1		1
18.	AQMA retrieval	6	1		7
19.	Reduce conversation with interactive retrieval	4		1	5
20.	English language retrieval			1	1
21.	Standards Report more frequently	2			2

retrieval should be by source type within a state, but they were not able to site specific applications. Among the Regional Offices only Region V and Region II indicated that the system might be useful to them. The other eight Regional Offices indicated that they would not use the system, or that any use would be minimal. Any applications for this system within the Regional Offices could be expected to take the place of current manual activities. Most Regional Offices have staff members contact states directly for update information, or they use the Bureau of National Affairs (BNA) publication for updates. In any case the consensus is that interactive retrieval is necessary before any advantage can be gained from the system. No estimates of the time spent manually updating regulation information was available in the survey.

Specific comments for this system follow.

- The major use of this system within the Regions will be to provide example texts for states that want to write new regulations.

- Interactive retrieval is necessary. Otherwise, current mechanisms for obtaining regulation information are sufficient.

- The system should contain state-approved as well as Federally approved regulations. The four steps in adopting a regulation as part of a SIP are: 1) State proposal of regulation; 2) State approval; 3) State request for EPA approval; 4) EPA approval.

- Retrieval by source type within state should be available.

- A report showing most current regulation numbers and changes within a specified period might be useful.

- The anticipated lag time for updating the data base makes the system less desirable for Regional Office and State use.

Not enough information is available at this time to translate these comments into specifications for making the system responsive to Regional Office needs.

2.2.6 APER Form Recommendations

The APER Forms were discussed only with the Regional Offices. The comments received indicate that modification of the forms and of the instructions associated with the forms is necessary. Changing the format of the forms is not warranted because of the time and cost involved. The implementation of a limited number of industry - specific forms may be the best approach.

The general comments received concerning APER forms are listed below.

- The generalized form is insufficient for many industries or process types, for example, coking, pulp and paper, smelters, phosphate fertilizers, refineries, and cotton ginning. Each of these operations requires specific information, moisture content of bark burned at paper mills, for example, not included on the APER forms. Since OMB clearance is required for changing existing forms or for adding new forms, the Regional Offices should be asked to document their industry specific requirements.

- The maximum percent sulfur should be designated as maximum percent burned during the previous twelve months and the maximum percent anticipated to be burned in the next twelve months. This information is useful only if the forms are completed, returned to the Regional Office, and reviewed within a few weeks. This is usually not the case. Normally, data may be one year old or even older before they reach the NEDS system. This application is not applicable to NEDS, but it is rather of interest for special studies using the APER forms. Inclusion of this information is a potential source of confusion for transferring data from APER forms onto NEDS coding sheets.

- A complete sketch of each process should be included as an attachment to the forms. All sketches should be labeled with the source codes used in the forms. This requirement should be printed on Page 1.

- The source code is probably the greatest problem with completing the forms. The instructions should require that one source code for each source be followed throughout the forms. The source code should also be reflected on the process diagram. The diagram should also make clear the stack and source relationships.
- A line should be added on Page 1 for the name of a corporate officer, and he should be required to sign the form. This requirement may help to improve the validity of the data.
- More explanation should be put at the top of each column rather than relying on the footnote technique. This may eliminate some problems associated with people ignoring the footnotes.
- On Page 2, a column should be added to allow identification of intermittent use of standby boilers.

2.2.7. QAMIS Recommendations

Current plans for QAMIS include the assignment of a Quality Control Index for each site, and the retrieval of quality control information by site, pollutant, laboratory, agency, or any combination of these four types. No further development is planned for this or any other system for quality control information until the Quality Assurance and Environmental Monitoring Laboratory (QAEML) provides guidelines. The Regional Offices would like the opportunity to comment on formal quality control requirements before requirements are promulgated.

The comments from the Regional Offices concerning the interim system indicate that any quality control grade assigned for a site should not be included on SAROAD reports. The concensus is that public criticism of monitoring activities or associated quality control procedures might impair Regional Office efforts toward implementing quality control programs among the states. This is especially true if the rating is subjective and does not take all factors into account.

Also, the judgment of data quality must be separate for each site/pollutant/method code. Selective judgment should be used for each site. Since quality control is a dynamic constantly changing discipline, it is probably that the information on many of the original "Data Quality Information Sheets for Air Pollution Agencies," which provided the data base for QAMIS are already outdated. Consequently, the best application for QAMIS at this time might be to indicate the relative status of quality control programs within each Region. This application would require frequent update of the original questionnaires by the affected agencies as their quality control procedures are implemented.

2.2.8. Administrative Recommendations

Several comments were made which relate to broadening the user base for the NADB systems. All of these comments were related to making the system capabilities more visible to users or potential users. Increased knowledge of and confidence in the systems would be expected to increase the number of users. Increasing the number of users could, in turn, result in an improvement in data quality as well as in an increase in suggestions for making the systems more responsive to user requirements.

The two most significant recommendations were:

- NADB should regularly publish system information in the "EPA Systems News" published by MIDSD. This activity could make more people outside of, as well as within, EPA more aware of data and systems capabilities available to them through NADB. Steps to implement this recommendation have already been taken.

- NADB should initiate an information distribution campaign among all control agencies and selected research and planning agencies who have a need for air quality or emissions data. Seminars could be offered to explain the operation of the systems and the applications for the data. Example report formats could be distributed along with an explanation of how each report might relate to specific activities.

These recommendations are supported by the fact that in this survey and in the non-EPA users survey ("Establishment of a Non-EPA User System for State Implementation Plans") the majority of people interviewed had little or no knowledge of the data bases maintained by NADB.

2.3 APPLICATIONS CATEGORIES

This section relates the survey results to three applications categories:

- Data base operations
- Data analysis
- Modeling The general benefits to each category that could be derived from implementation of the survey recommendations are discussed.

2.3.1 Data Base Operations

The major problems with system techniques and operating procedures for data input, storage, and retrieval in the NADB systems are:

- Currency of the data
- Data quality
- Responsiveness of the systems for report retrieval.

Table 7 shows the survey recommendations related to data base operations for each system. The general benefit from implementing any or all of these recommendations will be to increase the number of data users. The result of increasing the system usage is expected to be that more reliable data might be submitted if the organizations who submit the data also use the data base. This is not now the case. State agencies submit the bulk of the data to both NEDS and SAROAD, but they make little use of the systems.

Table 7. SURVEY RECOMMENDATIONS RELATED TO DATA BASE OPERATION

System	Application (see Tables 1 and 4)
NEDS	1, 5, 9, 10, 14, 15
SAROAD	3, 4, 10, 14, 19, 20

Most states are not aware of the retrieval and analysis capabilities of the NADB systems. Many of them have implemented their own automated systems. The same problem can be seen within the Regional Offices and within the NERC, RTP. At least three offices within the NERC have developed or are in the process of developing systems and data bases that derive their input from the same sources as the NADB data bases. The result may be that duplication of effort is diluting development resources that could be used to expand and improve the major NADB systems (NEDS and SAROAD). The next three sections summarize the recommendations for each of the three problem areas associated with data base operation.

2.3.1.1. Currency of the Data -

A major problem with keeping the data bases for NEDS and SAROAD current is that the data submittal and update procedures have inherent delays. Each state has forty-five (45) days after the end of the quarterly or semiannual reporting period cutoff to submit their reports. This means that, assuming a state holds all of its data until the end of a reporting period, the air quality data collected at the beginning of the quarter is almost five months old by the time it reaches the Regional Office; emissions data collected at the beginning of a semiannual reporting period can be almost eight months old by the time it reaches the Regional Office. Review procedures at the Regional Office and at NADB can add further delays of one to two months before update. In many instances, if the Regional Offices have had the data keypunched, further delays have been incurred.

Three approaches can be taken to avoid the delays at the state level:

- EPA can change the reporting requirements for both air quality data and emissions data. This approach is the least desirable because of the administrative considerations.
- EPA can provide a mechanism by which each state can enter update data to a central computer via a

remote terminal. The Regional Offices can then access the update data, perform validation and edit procedures, and release the data to NADB. An approach similar to this was explored in a previous report ("Establishment of a Non-EPA User System for State Implementation Plans", Contract No. 68-02-1001). Implementation of this approach solely for the purpose of expediting the update procedures would not be cost effective.

° The Regional Offices can work more closely with the states in finding a mechanism for allowing the states to submit data more frequently. This could be accomplished through installation of CDHS, or through providing assistance for interfacing each states air data systems with NEDS and SAROAD. Both approaches are being used. After a state's data processing system becomes operational, the generation of magnetic tapes in NEDS and SAROAD format directly from the state's system should be possible. Once NEDS and SAROAD compatibility are achieved, the cost of each update should be relatively small, and updates could be made more frequently than at present. This approach has the added advantage of avoiding keypunch bottlenecks at the Regional Offices. Moreover, the data edit and validation procedures could be expected to be faster, since preliminary edit and validation could be done at the state level. This approach should be formulated into a policy, since Regional program priorities differ.

NADB has already initiated steps to improve the update procedures between the Regional Offices and NADB. The planned decentralization of the edit and validation functions to the Regional Offices will eliminate one communication step currently necessary for updates. Implementation of a lowspeed terminal update capability at the Regional Offices will improve the time required for resubmittal of data that have been rejected by the edit. This capability can also allow Regional Offices to edit data that are collected sporadically as a result of special studies conducted by the Regional Offices or their contractors. This will also enhance the ability of the Regional Offices to comply with the requirements of "EPA Order No. 7520.2-Policy, Procedures, and Responsibilities for the Collection and Storage of Air Quality and Source/Emissions Data". Item 3 from Order No. 7520.2 is especially pertinent.

"... Policy - The Environmental Protection Agency will collect air quality and emissions-related data in accordance with the following:

a. Official EPA - and OMB-approved forms for ambient air data collection (SAROAD Form: OMB Number 158-R0012) and source/emissions data collection (NEDS Form: OMB Number 158-R0012) and Air Pollution Emission Report APER Form: OMB Number 158-R75) will be used by all EPA personnel involved in collecting, verifying, and updating such data.

b. Prior to initiation, all EPA projects, whether conducted in-house or by grant or contract, involving the collection of air quality or source/emissions data will be coordinated with the National Air Data Branch, Monitoring and Data Analysis Division, Office of Air Quality Planning and Standards, Research Triangle Park, North Carolina 27711.

c. All ambient air data and source/emissions data collected by EPA personnel or their representatives will be submitted to the National Air Data Branch in the proper format and in a timely fashion according to a schedule agreed upon at the beginning of the project...."

2.3.1.2 Data Quality -

Poor data quality is a criticism that is frequently offered by Regional Office personnel as an excuse for not accessing NEDS and SAROAD. The major problems with the data quality are related to misinterpretation of coding instructions, inadequate validation routines, and missing data.

Problems with coding NEDS are expected to diminish if NADB continues to offer user seminars. Most Regional Offices indicated that the seminars presented to date have provided significant assistance to people who prepare the data. Several Regional Offices also recommended that similar seminars for coding of SAROAD data be implemented. Suggestions concerning the structure of SAROAD seminars were not offered.

NADB has initiated the development of more validation routines for both NEDS and SAROAD data that will soon be available to the Regional Offices. Although the availability of more validation checks is expected to improve the overall data quality, if not handled properly it could negate the beneficial effects of decentralized editing and validation, which is intended to make the data bases more current. Verification and correction of data problems flagged during editing and validation procedures are manual tasks. Moreover, they usually require at least one communication between the Regional Offices and the states. Standard procedures should be implemented within the Regional Offices to insure that data that have passed edit and validation are forwarded to NADB immediately. These data should not be held until problems with other data in a batch have been corrected. Development of a bookkeeping system to trace dates, problems, and follow up for data flagged during editing or validation should be implemented. Currently, no mechanism exists for the Regional Office to easily determine the status of data that have been returned to the states for verification. An aging priority system that would indicate to each Regional Office the amounts and nature of data that have been held beyond a maximum time could improve the currency and the quality of the data bases. Neither NADB nor the Regional Offices have tight control over this situation.

Some potential users have indicated that they have not used the NADB systems because the data bases are missing significant data elements. A report is available to indicate for each state the total number of each parameter in NEDS that are missing. This report, however, gives no guidance on the priorities for obtaining missing data. Priorities should be set for the items most directly applicable to modeling or to fuel studies and trends analyses. The Regional Offices should use this report as a yardstick for determining NEDS update requirements of the states, and the states should be urged to provide these data with their semi-annual reports.

Determination of the degree to which the NEDS data base represents the total population of facilities also causes problems. NADB has defined formal procedures for verification of sources. The Regional Offices could use a set of internal guidelines for time limits and for methodology to be followed by the states in verification. NADB should also set time limits on the Regional Offices for verification response. Implementation of the CDS-NEDS cross-reference function should provide some assistance in verification.

Similar problems exist for SAROAD data. With SAROAD data, the problem could be caused by changing Agency/Project codes for a site or by data not being submitted for a sampling period. The problem with changing Agency/Project codes can be solved by expanding the Site Description Record and by maintaining the dates and numbers for changes at each site. Specifications for this recommendation are in Appendix B. The benefits to be derived, in terms of improved trends analyses capabilities and savings in manpower, offset the costs for this application, and it should receive high implementation priority. Implementation of an audit procedure to trace submittals for SIP required sites should also provide assistance for tracing data submittals for active sites.

2.3.1.3 Responsiveness of Systems for Report Retrieval -

Another reason given for not using the system is the lag time between requesting and receiving a report. Implementation of an aging priority system to assure rapid turnaround can help to solve this problem, and it will not impact the manpower at NADB or at the EPA Computer Center at Research Triangle Park. No changes to the NADB systems are required. This can help to assure a two day turnaround within the computer center. An approach should be explored with the computer center. Delays in transmitting batch reports from NADB to the Regional Offices constitute a separate problem. Remote batch printing on Regional Office printers has been mentioned by several Regional Offices as a significant problem. Backlogs on the Regional Office printers can cause delays as long as those normally anticipated by mailing reports. These delays negate any benefits derived from terminal batch processing. This is not a problem for NADB, but it is mentioned here because of its impact on the effectiveness of NADB procedures.

Implementation of an interactive file browsing capability would be expected to improve report turnaround time. Several users have commented that an indication of the completeness of parameters in the data base that pertain to a specific project could influence the decision to request a report. Not enough information is available to perform a cost/benefit analysis on this capability.

Implementation of data retrieval by class of parameter could improve report turnaround time. Total machine time for generating reports could be reduced. The extent to which parametric data retrieval will be implemented will depend on the file management system used.

2.3.2 Data Analysis

The major problems with data analysis using the NADB system are associated with:

- Data base discontinuity
- Definition of user requirements
- Quality control in data collection

Table 8 shows the survey recommendations related to data analysis for NEDS and SAROAD. The other systems are not a significant part of this discussion. The general benefit to be derived from implementation of any or all of these recommendations is the expansion of applications for NEDS and SAROAD. The number of users should expand as a result. Some duplication of system and program development efforts within EPA can probably be avoided.

Table 8. SURVEY RECOMMENDATIONS RELATED TO DATA ANALYSIS

System	Application (see Tables 1 and 4)
NEDS	1, 2, 3, 4, 7, 8, 11, 12, 13, 16, 17
SAROAD	1, 2, 6, 7, 8, 9, 11, 12, 14, 15, 16, 17, 18 19, 20, 21

2.3.2.1 Data Base Discontinuity -

A major problem affecting both NEDS and SAROAD is that update data are often not traceable to existing facilities in NEDS or sites in SAROAD. The NEDS problem has been caused by differences in Regional Office priorities combined with a lack of interest in NEDS among the states. Many states have made no attempt to maintain a cross-reference between sources in their permit or emission inventory systems and sources in NEDS. The result is that updates require manual cross checking of files at a point source level. Because of differences in defining point sources and/or because of the absence of common identifiers such as the equipment numbers used within a facility to identify individual pieces of equipment, cross checking NEDS point sources against state point sources appears to be almost impossible in many cases. The result is that the only way to update NEDS for some states, without huge manpower expenditures, is to replace the entire data base with a new inventory. Part of this problem is the result of states not being required to have significant ongoing NEDS coding instructions (APTD 1135) for proper cross-referencing of sources and source documents. The current implementation of expanded comments capabilities provides more opportunity for cross-referencing NEDS against state source documents. Although NADB has suggested using the Card 7 capability to cross-reference against state source documents, no formal requirement has been written. One reason often given for not using NEDS is the absence of any cross-reference information. NADB should consider incorporating the source document cross-reference requirement into APTD 1135. The capability to include cross-reference numbers has always been available in Card 6 of NEDS, but cross-reference numbers have not been included. Formal requirements for cross-reference numbers and incorporation of the requirements into the edit rejection criteria will eliminate the necessity for replacement of entire inventories in the future and result in a data base that can be used for emissions trends analysis. No impact on resources allocated for NEDS updates would be incurred.

Implementation of a CDS-NEDS cross-reference number can provide NEDS with still another source of information for updates. This capability requires that system development costs be incurred. In addition, it will have limited impact on the operating budgets of both NADB and the Division of Stationary Source Enforcement (DSSE). The benefits to be derived from this application can be directly related to improvement of the data base continuity, but a more important benefit is the improvement of data base quality.

SAROAD problems with data base continuity are primarily system problems, although associated operating problems are also apparent. Currently, if an agency code changes or if a modification at a site occurs, a new Site Identification Record is generated, and the old Site Identification Record is retained. Data for the new site are added to the summary files; however, the "old" site is carried in the Site Identification File, but no data for the old site code are added to the summary files. The only mechanism for tracing data, for a site at which site modifications or change in agency jurisdiction have occurred, is manual review of site records. This discontinuity of data that occurs for a site causes serious delays and problems for trends analysis. Recommendations have been made for flagging sites as active/inactive and for maintaining a history of changes to each site so that data can be easily traced. Users outside of EPA, and in some cases EPA users outside of NERC, RTP, may not be aware of the problems associated with the data continuity. Consequently, they risk reaching erroneous conclusions when performing trends analyses, or they may prematurely decide that not enough data for a geographical area exist in SAROAD to warrant analysis. Implementation of the recommendations mentioned above should improve the credibility of analyses performed on SAROAD data as well as expand the number of users.

2.3.2.2 Definition of User Requirements -

For both NEDS and SAROAD, better communications are needed between data base users and NADB regarding the kinds of analysis capabilities that are needed. One mechanism for improving communications is for NADB to publish system development plans in the "EPA Systems News." Duplication of efforts to develop systems within other EPA offices might be avoided. Specific requirement capabilities were defined for SAROAD trend analyses, but not for NEDS. One reason may be that the applications for emissions trends analyses are rather limited. The Office of Planning and Evaluation indicated that historical trends of emissions reductions by area and by source type are desirable. The NEDS is being modified to make year-of-record a key item. Any successful applications for emissions trends analyses will rely directly on solutions to the problems mentioned in Section 2.3.2.1. Most administrative offices could benefit from a trends analysis capability, but better problem definition will be required.

General requirements for trends analyses of SAROAD data have been defined. Most of the applications involve comparison of air quality figures quarterly or annually over several years against a baseline, usually an air quality standard. The capability to plot these data could save significant manpower within the Regional Offices. Successful implementation of trends analyses capabilities within SAROAD will depend on the solution of the data base problems.

Development of a general statistical package does not seem warranted at this time. Most users could be satisfied with Statistical Analysis System (SAS) package. Some operational problems exist with SAS, however, and these will need to be solved in order to satisfy current needs for statistical analyses.

2.3.2.3 Quality Control in Data Collection -

Quality control problems exist in the collection of both NEDS and SAROAD data. For both systems the implementation of more validation checks should improve the confidence of the data users. The major quality control problems that will still exist are those associated with data preparation.

For NEDS input, the quality control problem can be solved by following four recommendations; (1) all manual calculations should be shown on the back of any NEDS sheets submitted, (2) a specified percentage of calculations should be checked to ensure proper application of formulae, assumptions, and emission factors, (3) a specified percent of NEDS forms submitted should be checked against the source documents to ensure that data are being transferred properly, (4) identification of the data source should be coded into Card 7. These and other quality control procedures should be incorporated into APTD 1135. Guidelines should be developed to determine the portion of data collection budgets that should be allocated to data validation. This approach could be especially important in state agencies where several field offices participate in completing NEDS. Guidelines for quality control procedures in data handling should be developed by NADB, but final responsibility for implementation must be the responsibility of the Regional Offices.

Quality control criteria for SAROAD data are the responsibility of QAEML. There is currently no method for comparing the reliability of air quality data submitted by different agencies. Quality control techniques and auditing procedures should be uniform for all agencies submitting data to SAROAD. The only suggestion received during the survey regarding quality control data was for the incorporation of a quality control program scoring grade. Some Regional Offices object to this approach because of the possible public relations implications for some of their states. The incorporation of quality control into the collection of SAROAD data may require a separate quality control screening process that incorporates data from independent measurement system audits, calibrations, site visits, and laboratory quality control procedures. The current QAMIS system provides interim data regarding quality control procedures, but it is inadequate for indicating the effectiveness of those procedures.

2.3.3 Modeling

Most of the survey recommendations that can be related to modeling include applications to improve the data preparation steps. These recommendations are listed in Table 7.

In NEDS, implementation of a latitude - longitude to UTM conversion program, the polygon retrieval capability, and the confidentiality reporting option all have the potential for saving manpower. Implementation of the recommendations in Section 2.3.1.2 are especially applicable to modeling procedures, since they will help to assure that the data base is maintained in current status. Significant time and effort is expended collecting or validating emissions inventory data at the local agencies within states that are being modeled.

SAROAD applications for calculating wind roses and pollutant roses, for including data not meeting the 75% completeness criterion for computing averages, and for tracking data for a particular site all have the potential for expanding the effective data base available to users. As a result, users will have expanded bases for determining which air quality data more adequately describe an area. In many cases modelers currently must use data for the year that has the greatest volume of data, and no opportunity is available for comparing data collected for several years.

2.4 IMPLEMENTATION PRIORITIES

2.4.1 NEDS

2.4.1.1 Discussion of Benefits -

There are four identifiable benefits that may result from the implementation of the applications under consideration:

- An improvement of the quality and completeness of data in NEDS.
- Reduction of input/output time and turn around time.

- A net savings of man-hours.
- An improvement in usage of NEDS.

Benefits to be derived from the 17 applications under consideration are summarized in Table 9.

It is not feasible to place a dollar value on benefits, since usage statistics for various applications were not available in the survey. As has been stated previously, improvements in data quality and the response time of NEDS are incentives to make better usage of the system. More usage of NEDS should result in a higher quality of data submitted by the State and local agencies.

A total of five (5) applications are expected to result in a direct cost savings. Each application and the expected savings is discussed below.

- Latitude-longitude with UTM Conversion - Currently, many of the agencies submitting NEDS data use maps prepared by their respective Highway Departments. These maps are typically marked with only latitude and longitude. In such instances, rather than purchase the necessary USGS maps with UTM markings, the agency submits their NEDS data without grid coordinates. To determine the dollar cost of this application it is necessary to determine the number of agencies that can provide only latitude-longitude coordinates and the cost of purchasing the necessary USGS maps.

- Retrieve by Range of Parameter - This application will save computer time primarily through a reduction in the lines of print associated with reports that list individual point sources. Because usage statistics for the various NEDS reports are not available, it is difficult to determine the dollar savings with any degree of precision. Users of the data can also expect to reduce the man-hours spent in reviewing the now lengthy data listing.

Table 9. BENEFITS TO BE REALIZED FROM IMPLEMENTATION OF APPLICATIONS

NEDS

No.	Application	Benefit			
		Improve data quality	Reduce input/output time	Save man-hours	Improve usage of NEDS
1.	Latitude-longitude with UTM conversion	X		X	
2.	Polygon retrieval		X		X
3.	Retrieve by range of parameter		X	X	X
4.	Analyze effect of potential changes to parameters in source record			X	X
5.	Lowspeed terminal update	X	X		
6.	Usage statistics by report and/or user				X
7.	Basic accounting capability				X
8.	Emission trends			X	X
9.	Decentralization of edit	X	X		
10.	NEDS/CDS cross-reference	X			X
11.	Expanded comments capability	X			X
12.	Indicate sources deleted from reports because of confidentiality			X	X
13.	AQMA retrieval	a			
14.	Implement user training seminars	X			X

[a] No significant benefit could be determined.

Table 9 (continued). BENEFITS TO BE REALIZED FROM IMPLEMENTATION OF APPLICATIONS

NEDS

No.	Application	Benefit			
		Improve data quality	Reduce input/output time	Save man-hours	Improve usage of NEDS
15	Discontinue or modify verification report	X			
16	Add fugitive dust SCC's	X			
17	Improve the format for identifying report retrieval options.				X

° Analyze Effect of Potential Changes to Parameters in Source Record - This application will result in a direct savings of man-hours when it is necessary to calculate potential emissions from a large number of point sources in a geographical area. Suppose an area has 200 point sources. Working from a NEDS point source listing one could expect than an average of 5 minutes/point source or approximately 16 man-hours would be required to do the necessary calculations.

The computer time required to calculate potential emissions is not expected to exceed the computer time required to obtain the point source listing required for the manual calculations.

- Emission Trends - If this application is made available to the users it is reasonable to assume that emission trends would be determined on an annual basis for perhaps 300 geographical areas. To manually determine the trend for any given area may require 2 manhours. This time would be spent preparing the necessary request for data retrieval from NEDS and the calculations required to fit a trend curve.

- Indicate Sources Deleted from Reports Because of Confidentiality - Users must now review all reports to determine if there are sources, classified as confidential, that have not been included in various totals. Depending upon the specific report format and the number of point sources in the area such review may take 1-2 manhours per report.

2.4.1.2 Assignment of Priorities -

NEDS users generally agree that the implementation of programs to improve data quality and procedural changes to improve data flow should receive the highest priorities by NADB. The implementation of these programs and procedural changes will not require major changes to NEDS.

The implementation of programs to improve data quality will require a commitment of additional effort on the part of the State and local agencies in preparing NEDS data for new and modified sources of emissions. Regional Offices will have to expedite procedures to improve data flow to NADB. Finally, NADB must initiate action to assure minimal turnaround time with the computer center.

A total of thirteen (13) applications were identified which require the development of new programs or revisions to existing programs in the system. Four (4) applications were identified which require changes in current operating procedures. The PEDCo project team assigned implementation priorities for each application (Table 10). The priority assignments are based on a subjective assessment of the importance of each application as discussed with the survey participants. The highest priorities are assigned to applications 9, 5, 14, and 11, none of which are related to the actual user applications for the data. These applications all relate to either improving data quality (9 and 11), getting data into the system more rapidly (5), or increasing the number of users (14).

2.4.2 SAROAD

2.4.2.1 Discussion of Benefits -

Benefits to be derived from the 21 applications being considered in this survey are summarized in Table 11. Twelve applications have multiple benefits. For example the use of lowspeed terminal update is expected to improve data quality, reduce input/output time and save man-hours on the part of the system users. It is interesting to note that sixteen (16) of the applications are expected to improve usage of the system. Based upon the information provided by the users contacted as a part of this survey, eight (8) applications should result in a direct savings of man-hours. Each application is discussed below.

- Trends Plotting - To manually determine and plot quarterly trends requires about 1 man-hour per pollutant for each station. Considering only those stations required by the SIP's and the pollutants required to be monitored, there are 3,024 station-pollutant combinations. On an annual basis, then 3,024 man-hours might be saved by the use of an automated trend plotting procedure.

- Audit SIP Station Reporting - Regional Offices spend a significant amount of time tracking the quarterly air quality data submitted by the States.

Table 10. PRIORITY ASSIGNMENTS FOR NEDS APPLICATIONS

No.	Application	Priority
1.	Latitude-longitude with UTM conversion	10
2.	Polygon retrieval	15
3.	Retrieve by range of parameter	13
4.	Analyze effect of potential changes to parameters in source record	9
5.	Lowspeed terminal update	2
6.	Usage statistics by report and/or user	16
7.	Basic accounting capability	14
8.	Emission trends	11
9.	Decentralization of edit	1
10.	NEDS/CDS cross-reference	8
11.	Expanded comments capability	4
12.	Indicate sources deleted from report because of confidentiality	5
13.	AQMA retrieval	17
14.	Implement user training seminars	3
15.	Discontinue or modify verification report	6
16.	Add fugitive dust SCC's	12
17.	Improve the format for identifying report retrieval options	7

Table 11. BENEFITS TO BE REALIZED FROM IMPLEMENTATION OF APPLICATIONS

SAROAD

No.	Application	Benefit: Improve data quality	Benefit: Reduce input/output time	Benefit: Save man-hours	Benefit: Improve usage of NEDS
1.	Polygon retrieval		X		X
2.	Trends plotting			X	X
3.	Audit SIP station reporting	X		X	
4.	Lowspeed terminal update	X	X		
5.	Usage statistics by report and/or user				X
6.	Graphics to plot site locations				X
7.	Calculate wind rose and pollutant rose			X	X
8.	Parametric data retrieval		X	X	
9.	Report on standards violations				X
10.	Decentralized edit	X	X		
11.	Option to include data not meeting 75% criteria				X
12.	Include 2nd maximum on Standards Report			X	X
13.	Identify inactive sites		X	X	X
14.	Include reporting units on edit and validation reports	X			X
15.	County retrieval		X	X	X
16.	Include minimum detectable level on report				X

Table 11.(continued) BENEFITS TO BE REALIZED

IMPLEMENTATION OF APPLICATION

SAROAD

	Application	Benefit			
No.	Application	Improve data quality	Reduce input/output time	Save man-hours	Improve usage of SAROAD
17.	Include accuracy of method in site I.D. file				X
18.	AQMA retrieval	a			
19.	Reduce conversation with interactive retrieval		X	X	
20.	English language retrieval				X
21.	Standards Report more frequently				X

[a] No significant benefit could be determined.

It is estimated that such auditing requires about 30 man-hours per state per quarter. On an annual basis this amounts to more than 2500 man-hours for all Regions combined. Even with an automated audit procedure some manual effort would still be necessary. Assuming the net savings of man-hours is only 1250 hours, the annual savings may still be as much as $7,500.00 based on a cost of $6.00/hr. This application also has the added benefit that the quality of data in SAROAD will be improved significantly by a better audit mechanism.

- Calculate Wind Rose and Pollutant Rose - This application will save man-hours and also improve usage of the system. Based upon information obtained during this survey it is not possible to quantify the savings in man-hours. A feasibility study, however, has been made.

- Parametric Data Retrieval - This application has two possible benefits. The reduction in input/output time will result in a savings of the computer time now required to print lengthy reports. Users of SAROAD data can expect a significant savings in the man-hours required to retrieve the specific data needed for their own reports.

- Include 2nd Maximum on the Standard Report - With the present format of the Standards Report the user is required to look at a detailed tabulation to determine the 2nd maximum whenever the maximum exceeds the pertinent air quality standard. Certainly this is a measurable amount of time required to look up the 2nd maximum. Because the cost of implementing this change is expected to be negligible, The major benefit becomes improved usage of the SAROAD data base.

- Identify Inactive Sites - This application is expected to significantly reduce the volume of lines of print required on data tabulations and reports. Such a reduction will result in a direct savings of computer time. Users will realize a savings in man-hours by handling a lesser amount of paper. Finally, the ability to retrieve more specific data will be an added incentive for the users to access SAROAD. A feasibility study is needed to estimate the actual savings in man-hours.

- Retrieval by County - As seen in Table 15 this application can be expected to provide multiple benefits. Essentially the application allows the user to retrieve more specific data, which reduces the volume of outputs. The major benefit again is expected to be the incentive on the part of users to make greater use of the data base.

- Reduce Conversation with Interactive Retrieval - This application will result in a savings of telephone line charges, computer time and man-hours. An evaluation of the actual savings cannot be made without usage statistics. A feasibility study is required to evaluate this application.

2.4.2.2 Assignment of Priorities -

Users of SAROAD indicated that, for the most part, existing data retrieval formats provide the type of data listings and summary reports needed for planning, enforcement, and management decision making. As was previously discussed for NEDS, SAROAD users gave the highest priority to those applications that tend to improve getting data into, and out of, the system.

A summary of the priority assignments for the 21 applications is presented in Table 16 . A priority assignment for each application was made by the PEDCo project team.

The priority assignments made by the PEDCo project team are based upon a subjective assessment of the importance of each application, following the interviews of the various users of SAROAD. The highest priorities are given to: decentralized edit (1); audit of SIP reporting stations (2);and lowspeed terminal update (3). These applications are all concerned with reducing the time required to get new data into the data base. The identification of inactive sites is given a priority

Table 12. PRIORITY ASSIGNMENTS FOR SAROAD APPLICATIONS

No.	Application	Priority
1.	Polygon retrieval	17
2.	Trends plotting	9
3.	Audit SIP station reporting	3
4.	Lowspeed terminal update	2
5.	Usage statistics by report and/or user	16
6.	Graphics to plot site locations	15
7.	Calculate wind rose and pollutant rose	14
8.	Parametric data retrieval	12
9.	Report on standards violations	5
10.	Decentralized edit	1
11.	Option to include data not meeting 75% criteria	8
12.	Include 2nd maximum on Standards Report	11
13.	Identify inactive sites	4
14.	Include reporting units on edit and validation reports	13
15.	County retrieval	7
16.	Include minimum detectable level on report	18
17.	Include accuracy of method in site I.D. file	19
18.	AQMA retrieval	21
19.	Reduce conversation with interactive retrieval	10
20.	English language retrieval	20
21.	Standards Report more frequently	6

of 4, reflecting the need to reduce time spent in extracting necessary data from otherwise voluminous tabulations. The next ten applications have to do with more specific and more frequent standards reports.

2.5 IMPLEMENTATION REQUIREMENTS

This section summarizes the level of effort and cost for implementing each survey recommendation. The recommendations are divided into two classifications:

- Recommendations for improving data base operations.
- Recommendations for improving data analysis capabilities.

For recommendations specified in the Appendix, the level of effort and cost derivations are included with the specifications. Costs are estimated from:

- Required manpower-professional and clerical support (Clerical support includes keypunching)
- Compile and run times based on UNIVAC 1100 run times for similar programs
- Travel, printing, or other direct charges.

Manpower costs are based on the number of hours required for previous applications similar in type and scope to these recommendations. The cost of a professional man-hour is assumed to be $20 per hour, or an average rate that might be expected by using a mix of GS11 and GS12 levels. The average clerical support cost is assumed to be $5 per hour.

Program compilation and run times are based on UNIVAC times for current programs similar to the ones specifed. Computer charges are based on total estimated computer time charged at $368.00/SUP (SUP is System Unit of Processing). The SUP is a total of CC/ER charge, CAU Time, and I/O Time, where:

CC/ER = computer execution requests
CAU Time = time the central processor is active
I/O Time = number of file accesses.

It should be emphasized that run times are only estimates, and that for most applications they are extrapolated from run times for programs similar to but not the same as the specific programs discussed. Annual operating costs are not projected here because of the lack of available information from users concerning the number of job requests they have for current NADB programs as well as the requests they might have for projected programs(this supports the need for system usage statistics).

Travel, printing, or other direct charges involved with implementation or operation of any of the recommendations are based on current commercial rates.

2.5.1 Implementation Requirements for NEDS

The level of effort and cost for implementing each of the recommendations related to improving the data base operation are summarized in Table 13. The level of effort and cost for implementing each of the recommendations related to improving the NEDS data analysis capabilities are summarized in Table 14. For this application category, all annual operating costs would be in addition to current system operating costs. Where cost savings could be realized, they would be in terms of manpower costs versus computer costs.

A significant cost saving might be realized by combining the implementation efforts for:

- Retrieval by Range of Parameter
- Analysis of the Effects of Potential Changes.

If the NADB*NEDS USER file is defined as a system 2000 file that meets the requirements for both applications, then the costs anticipated for building the data base need to be incurred only once. As shown in Appendix A-3 and A-4, the System 2000 data bases that might be defined for each of the two applications are similar enough for this approach to be taken. Moreover, this approach might preclude the necessity for writing a separate program to fulfill the requirement for a basic accounting capability.

2.5.2 Implementation Requirements for SAROAD

The level of effort and cost for implementing each of the recommendations related to improving the data base operation

Table 13. NEDS DATA BASE OPERATION - IMPLEMENTATION REQUIREMENTS

No.	Application	Implementation requirements				Total
		Man-hours	Manpower costs	Computer costs	Other	
1.	Latitude-longitude	80	$ 1,360.00	$185.00		$ 1,545.00
5.	Lowspeed terminal update					5,000.00[a]
9.	Decentralization of edit					15,000.00[b]
10.	NEDS-CDS cross-reference	428	7,530.00	460.00		7,990.00
14.	Implement user training seminars					
	1) Data preparation	440/yr	7,600.00		$2,410.00/yr	10,010.00/yr
	2) User-oriented					
	a. Development (1-time cost)	680	11,800.00		1,250.00	13,050.00
	b. Presentation	440/yr	7,600.00		2,410.00/yr	10,010.00/yr
15.	Discontinue or modify verification report	(Total cost not estimated in Appendix, but expected to be less than $1,000.00).				

[a] Estimates for both NEDS and SAROAD indicate $10,000.00 for the total project.

[b] Estimates are that both NEDS and SAROAD changes will cost between $25,000 - $30,000.

Table 14. NEDS DATA ANALYSIS - IMPLEMENTATION REQUIREMENTS

No.	Application	Implementation requirements				Total
		Man-hours	Manpower costs	Computer costs	Other	
1.	Latitude-longitude input	80	$1,360.00	$185.00		$1,545.00
2.	Polygon retrieval	1274	23,920.00	1,475.00		25,395.00
3.	Retrieve by range of parameter	392	7,120.00	1,215.00		8,335.00
4.	Analyze effect of potential changes to parameters in source record	420	7,500.00	1,015.00		8,515.00
7.	Basic accounting	260	4,840.00	1,105.00		5,945.00
8.	Emission trends		(Not enough information to estimate)			
11.	Expanded comments capability					
12.	Indicate sources deleted from report because of confidentiality	67	1,295.00	185.00		1,480/program
13.	AQMA retrieval					
16.	Fugitive dust SCC's	666	11,520.00/SCC			11,520.00/SCC
17.	Improve the format for identifying report retrieval options	80	1,000.00		$500.00	1,500.00

for SAROAD are summarized in Table 15.

The level of effort and cost for implementing each recommendation related to improving SAROAD data analysis capabilities are shown in Table 16.

Some of recommended applications for SAROAD can be combined in the implementation phase, and a significant cost savings can be realized. The possibility of combining applications in the implementation phase is expected to affect the system of priorities applied to individual applications. The suggested areas for consolidating implementation efforts follow.

- Polygon Retrieval and Graphics for Site Locations - The program for polygon retrieval includes logic for determining if a point is within a defined area. The same logic is used in the program to provide graphic display of site locations. As much as 50 percent of the cost of implementing the latter capability might be saved by combining implementation.

- Polygon Retrieval for SAROAD is the same as for NEDS. Changes to the input/output formats and to the JCL is the major difference between these applications. At least 75 percent of the cost of this program for SAROAD should be saved if it is already implemented for NEDS, and vice versa.

- Several report programs will be affected if any of the following program changes are implemented:
    1. County retrieval
    2. Addition of an inactive site code
    3. Minimum detectable levels on reports
    4. Statement on method accuracy in site record.

Any schedule for implementation should reflect the fact that a significant amount of testing time, both programmer time and computer time, can be saved by combining the applications rather than making each change individually.

Table 15. SAROAD DATA BASE OPERATION - IMPLEMENTATION REQUIREMENTS

No.	Application	Implementation requirements				Total
		Man-hours	Manpower costs	Computer costs	Other	
1.	Audit SIP station reporting	288	$5,220.00	$735.00		$5,955.00
4.	Lowspeed terminal update					5,000.00[a]
10.	Decentralized edit					15,000.00[b]
14.	Include reporting units on edit and validation reports	(Total cost not estimated in Appendix, but excepted to be less than $1,000.00)				
19.	Reduce conversation with interactive retrieval[c]					

[a] Estimates for both NEDS and SAROAD indicate $10,000 for the total project.

[b] Estimates are that both NEDS and SAROAD changes will cost between $25,000 - $30,000.

[c] Requires a feasibility study.

Table 16. SAROAD DATA ANALYSIS - IMPLEMENTATION REQUIREMENTS

No.	Application	Man-hours	Manpower costs	Computer cost	Other	Total
1	Polygon retrieval	(See Table 14)				
2	Trends plotting	472	$8,360.00	$1,205.00		$9,565.00
6	Graphics to plot site locations	272	4,960.00	1,500.00		6,460.00
7	Calculate wind rose/pollutant rose[a]					30,000.00
8	Parametric data retrieval[b]					
9	Report of Standards Violations	160	2,840.00	275.00		3,115.00
11	Option to include data not meeting 75% criteria	144	2,640.00	285.00		2,915.00
12	Include 2nd maximum on Standards Report	86	1,600.00	185.00		1,785.00
13	Identify inactive sites[c]					
15	County retrieval	112	2,120.00	185.00		2,305.00 per program
16	Include mimimum detectable level on report	56	1,060.00	370.00		1,430.00
17	Include accuracy of method in site I.D. file[c]					
18	AQMA retrieval[c]					
19	Reduce conversation[c] with interactive retrieval					
20	English language[c] retrieval					
21	Standards report more frequently	(Implementation costs not required)				

[a] These figures were taken from a feasibility study on this problem.

[b] Cost for this may be as high as $50,000, depending on demonstrated capability of System 2000.

[c] Estimates not made. Feasibility study required to supply cost estimates.

APPENDIX A

TASK DESCRIPTIONS

NEDS

APPENDIX A TASK DESCRIPTIONS - NEDS

A.1.

APPLICATION - Latitude/longitude conversion to UTM coordinates in NEDS format.

ABSTRACT - The purpose of this program is to allow users to input latitude/longitude coordinates for conversion to UTM coordinates or vice versa. TCLCONV (FORTRAN IV) is already available to provide the conversion algorithms. TCLCONV will be modified for this application. The modifications will provide two types of output:

1. Punched cards with UTM coordinates in NEDS format for direct updating of NEDS.
2. An interactive listing of latitude/longitude coordinates and their corresponding UTM coordinates. This listing includes user comments, such as plant I.D.'s or sampling site I.D.'s, if the user provides them as input.

The first type of output should result in a more complete data base by allowing states to input UTM coordinates to NEDS when only latitude/longitude coordinates are available from the state files. The second type of output will be useful for such future applications as polygon retrieval (A.2). This application will be executed at the Regional Office or State Agency level to minimize the impact on NADB's operating procedures and to provide maximum turnaround efficiency.

INPUT DESCRIPTION - Input is punched card or keyboard input to TCLCONV. The input must include NEDS "key" identifiers if NEDS update cards are desired, user comments can be input along with the latitude/longitude or UTM coordinates. An example input description for generating NEDS output cards follows:

CC	Description	Format
1-2	SAROAD State No.	F2.0
3-6	SAROAD County No.	F4.0
7-9	Federal AQCR No.	F3.0
10-13	Plant I.D. No.	F4.0
14-15	Point I.D. No.	F2.0
16-20	Blank	5X
21-26	Latitude	F6.0
27-33	Longitude	F7.0
34-80	Blank	

An example input description for generating the interactive coordinates listing follows.

CC	Description	Format
1-20	User comments	A20
21-26	Latitude	F6.0
27-33	Longitude	F7.0
34-80	Blank	

This format will be changed slightly to accomodate the case in which UTM's are input.

OUTPUT DESCRIPTION - Output for the NEDS option is punched card from TCLCONV in NEDS format as "changes". A listing of input/output is printed.

Card 1 - One card per plant

CC	Description	Format
1-2	SAROAD State No.	F2.0
3-6	SAROAD County No.	F4.0
7-9	Federal AQCR	F3.0
10-13	Plant I.D. No.	F4.0
14-17	blank	5X
18-19	UTM Zone	F2.0
20-77	blank	58X
78	Action Code "C"	A1
79	"P"	A1
80	Card No. "1"	I1

Card 2 - One card per point source

CC	Description	Format
1-2	SAROAD State No.	F2.0
3-6	SAROAD County No.	F4.0
7-9	Federal AQCR	F3.0
10-13	Plant I.D. No.	F4.0
14-15	Point I.D. No.	F2.0
16-23	Blank	8X
24-27	Horizontal UTM	F4.1
28-32	Vertical UTM	F5.1
33-77	Blank	45X
78	Action Code "C"	A1
79	"P"	A1
80	Card No. "2"	I1

INPUT/OUTPUT LISTING FORMAT

STATE	COUNTY	AQCR	PLANT	POINT	LAT.	LONG.
XX	XXXX	XXX	XXXX	XX	XX XX XX	XXX XX XX

NORTHING	EASTING	ZONE
XXX.X	XXXX.X	XX

OPERATION - Input is either keyboard or punched cards in the format required by the output option the user specifies. If the NEDS output cards are desired, the complete NEDS key is mandatory. A control card can be used to identify the input/output options desired.

PROGRAMMING REQUIREMENTS - The following changes to TCLCONV are needed.

1. Change input formats to accomodate NEDS key data.
2. Change input formats to accomodate user comments.
3. Input a control card to identify the desired input conversion and the output option.

4. Define a card punch as an output device.
5. Format the output for each option.
6. Insert logic to punch a NEDS Card 1 for each plant if NEDS update cards are desired.
7. Insert logic to punch a NEDS Card 2 for each input card.

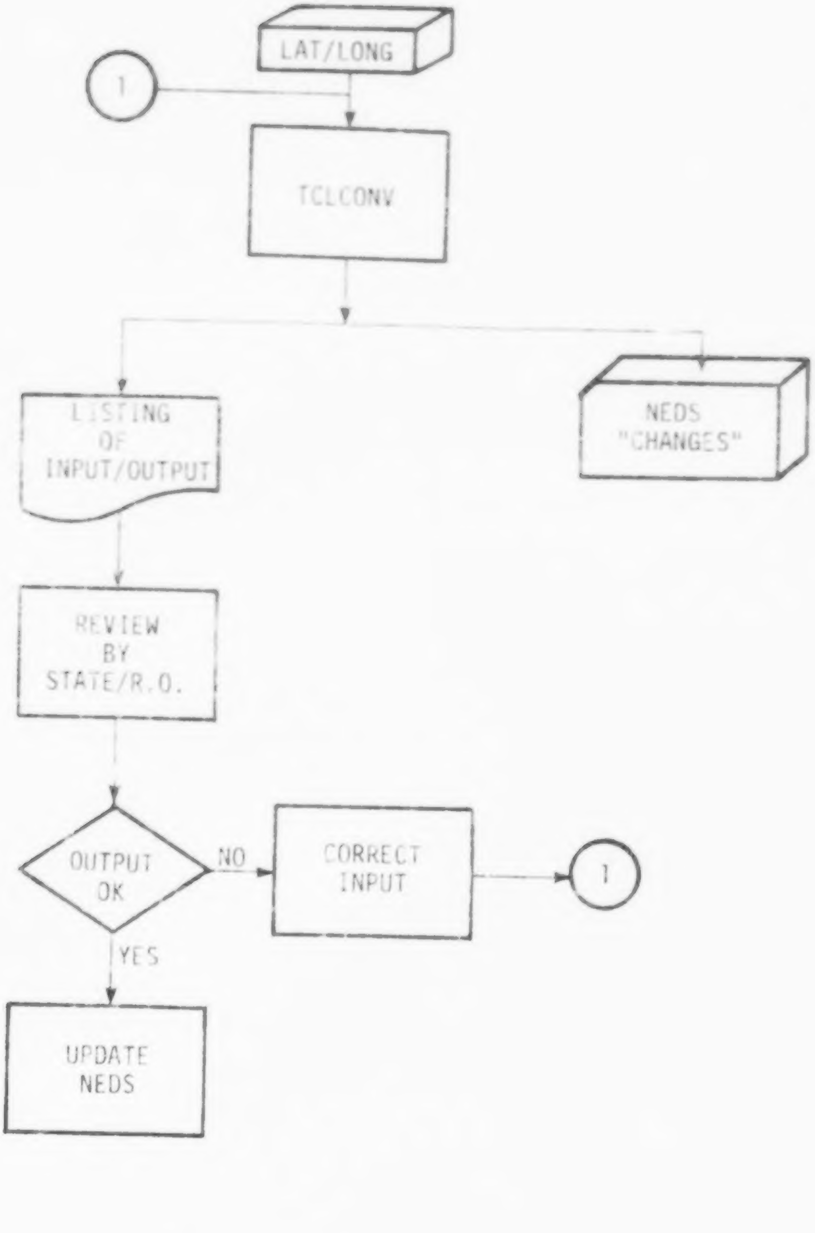
The flow diagrams for the logic to produce NEDS update cards are shown on the following pages.

LEVEL OF EFFORT and ANTICIPATED COST*-

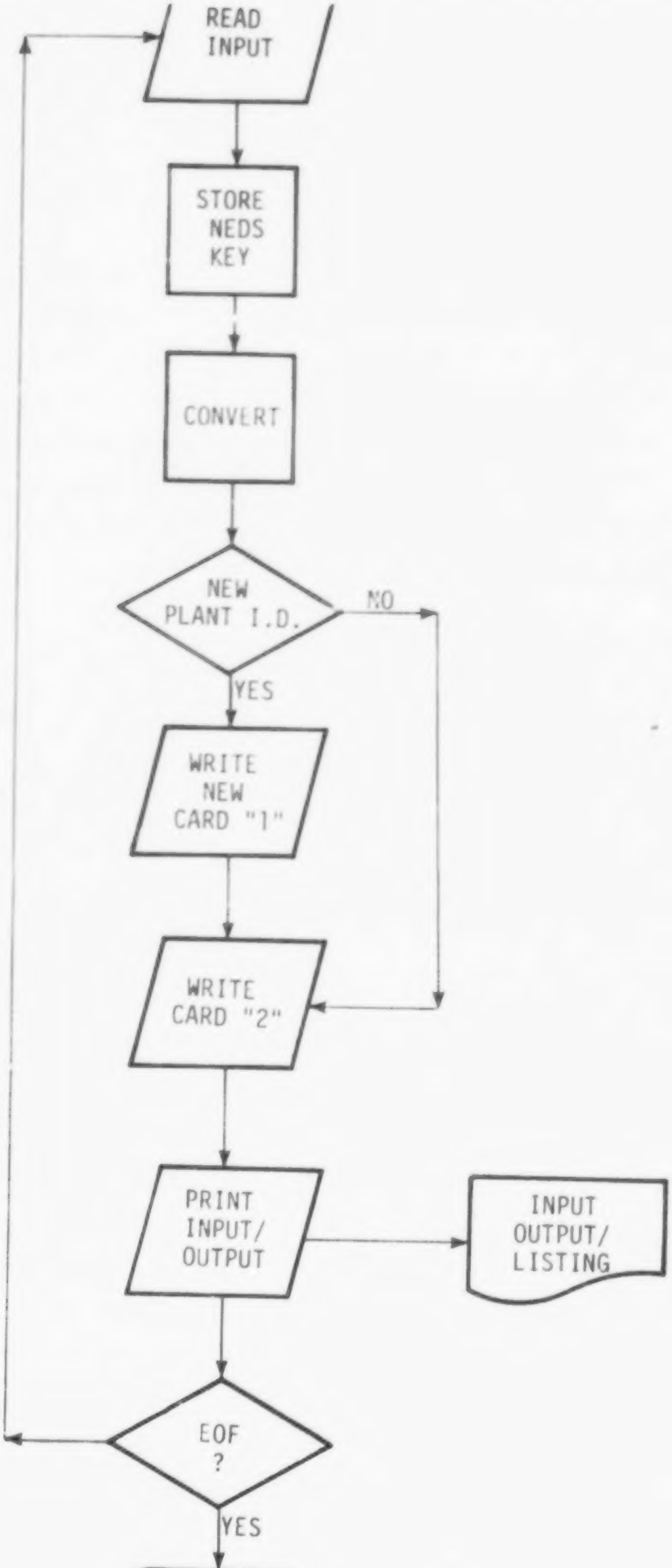
1.	Changes to TCLCONV = 40 hours @ $20/hr.	=	$ 800.00
2.	Documentation including instructions for R.O.'s = 24 hours @ $20/hr plus 16 hours clerical support @ $5/hr.	=	560.00
			1360.00
3.	Computer time is:		
	2 test @ 15 min/test including compile and run time = 30 min = .5 SUP @ $368/SUP	=	185.00
			$1545.00

* Note: Costs presented in appendices A and B are given to nearest $5.00.

SYSTEM FLOW


LAT/LONG
1
TCLCONV
LISTING OF INPUT/OUTPUT
NEDS "CHANGES"
REVIEW BY STATE/R.O.
OUTPUT OK
NO
CORRECT INPUT
1
YES
UPDATE NEDS

PROGRAM LOGIC


READ INPUT
STORE NEDS KEY
CONVERT
NEW PLANT I.D.
NO
YES
WRITE NEW CARD "1"
WRITE CARD "2"
PRINT INPUT/ OUTPUT
INPUT OUTPUT/ LISTING
EOF ?
YES
END

A.2.

APPLICATION - Polygon - defined area retrieval

ABSTRACT - The purpose of this program is to allow users to define, by latitude/longitude coordinates, a polygon-shaped geographical area and to retrieve data on NEDS point sources or SAROAD monitoring sites within the defined area. Two algorithms will be applied. The first is for determining if a point lies within a defined area. This algorithm must be developed. The second is for determining if a point lies within a specified radius of another point in the polygon. This algorithm has been developed but it requires slight modifications. These algorithms will allow users to:

1. Input latitude/longitude coordinates to NEDS and create a subfile of point sources within the polygon that meet selection criteria. This will create a subfile in the same format as NE001A.

2. Retrieve all site identification numbers for SAROAD monitoring sites within a radius of a specified point source.

3. Input coordinates to SAROAD and obtain an active site listing for monitoring sites within the polygon.

The rationale of the total capability described is to reduce print time and report size as well as to increase total report retrieval flexibility. This application will be in FORTRAN IV. Latitude/longitude coordinates are recommended instead of UTM coordinates, because the UTM coordinate system is based on zones around the globe, and there is a discontinuity between zones. Consequently, when a polygon area spans two zones, it is necessary to project coordinates from one zone into the other.

INPUT DESCRIPTION - Input will be punched cards or keyboard. A control card will designate the option desired. The options will be:

1. Create the subfile of NEDS point sources that meet a specified set of criteria and lie within the polygon.
2. Retrieve a list of NEDS plant I.D.'s that lie within the polygon.
3. Retrieve a list of SAROAD active monitoring sites for a specified pollutant/method that lie within the polygon.
4. Retrieve I.D. numbers for all point sources or monitoring sites that lie within a designated radius of a given NEDS point source.
5. For all NEDS point sources that meet specified criteria and lie within the polygon, create a subfile of I.D. numbers of SAROAD monitoring sites that lie within a specified radius of any of the point sources.[a]

---

[a] Before applying the radius selection algorithm, the SAROAD monitoring site I.D. should be compared to the I.D.'s already selected. This will save processing time and insure that each monitoring site is represented only once in the subfile being created. This test is necessary, since it is possible for a single monitoring site to be within the designated radius from multiple NEDS point source locations.

For each processing option, the control card format will differ, since different parameters must be specified. Separate formats for designating the political subdivisions will be applied for both NEDS and SAROAD.

NEDS POLYGON INPUT - For NEDS, the State, AQCR, and County code numbers can be input along with the latitude/longitude coordinates that define the area. A parameter "n" indicates the number of vertices (points) on the polygon. If County code is not entered, control is on AQCR code within State code. If County code and AQCR code are not entered, control defaults to State code. Two sets of cards are input. The first card identifies the NEDS keys. The second card or set of cards define the polygon. Multiple states can be run, but multiple request cards are required.

Card 1

CC	Description	Format
1-2	State I.D.	I2
3-6	County I.D. 1	I4
7-9	AOCR I.D. for County 1	I3
. . .	(Repeat County I.D. for up to 11 counties per card)	
78-79	Blank	X
80	"1"	I1

Repeat Card 1 for each state

Card 2

1-2	Number of points on the polygon	I2
3-9	Longitude coordinate for first point.	F7.0
10-15	Latitude coordinate for first point	F6.0

16-22	Longitude coordinate for second point	F7.0
23-28	Latitude coordinate for second point	F6.0
. . .	Repeat above format for each polygon point up to 5 points per card	
80	"2"	I1

Repeat coordinate fields on Card 2's until "n" fields equal to the number of points defined in the first Card 2 have been defined.

Polygon coordinate fields must be defined in sequence clockwise beginning with the point that has the largest longitude coordinate.

Card 3

1-4	SIC	I4
5-6	IPP	I2
7-14	SCC	I8
15-21	Operating rate	I7
22-24	Sulfur content	F3.1
25	Source code	A1
26	Run option	I1
27-28	Radius retrieval (km)	I2

This format could be altered to include any criteria selection parameters desired.

SAROAD POLYGON INPUT - For SAROAD, the State, AQCR, and Pollutant/Method codes can be input along with the latitude/longitude coordinates that define the area. A parameter "n" indicates the number of points on the polygon. If Pollutant/Method codes are not input, selection is by AQCR code within State code. If Pollutant/Method codes and AQCR code are not input, selection defaults to State code. The second card or set of cards define the polygon.

Card 1

CC	Description	Format
1-2	State I.D.	I2
3-5	AQCR	I3
. . .	(Repeat AQCR's for as many state portions of AQCR's as desired	
80	"1"	I1

Repeat Card 1 for each state

Card 2 - Card 2 format to identify the polygon is the same as for NEDS Card 2.

Card 3

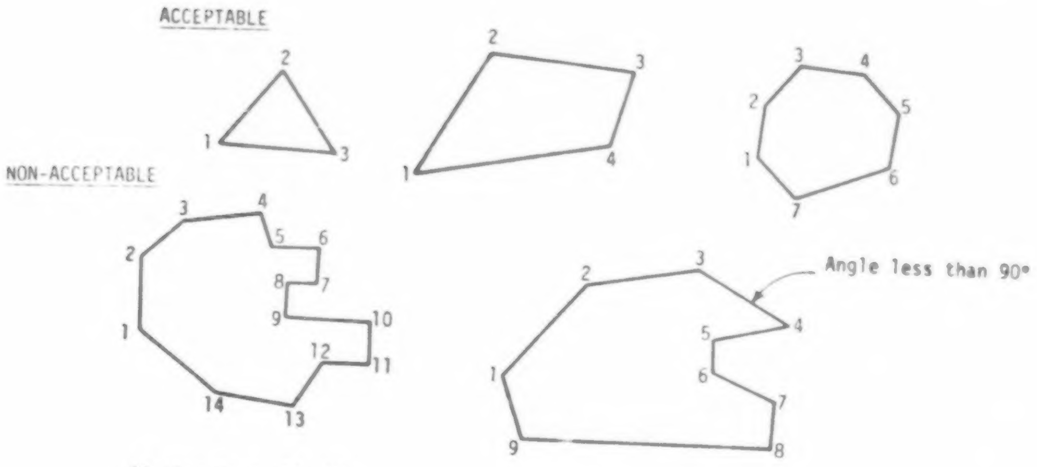
CC	Description	Format
1-7	Pollutant/Method code	I7
8	Interval code	I1
9-10	Year of record	I2

Repeat the above format for each pollutant/method/interval/year combination.

OUTPUT DESCRIPTION - Output will be the subfiles for application of report programs or an interactive listing of desired information. An example of an interactive listing is shown below for NEDS plants within a defined polygon.

```
                PLANTS WITHIN POLYGON DEFINED BY
            LATITUDE                      LONGITUDE
           XX  XX  XX                   XXX  XX  XX
           "   "   "                     "   "   "

STATE               AQCR            COUNTY          PLANT
  XX                XXX              XXXX            XXXX
  "                  "                "               "
```

OPERATION - Polygon Retrieval Algorithm - Input for defining the polygon must be sequential clockwise beginning with the point with the largest longitude coordinate.
The polygon can be any triangle, any rectangle, or any convex polygon, i.e. one for which the angle between any two consecutive sides is greater than 90 degrees but less than 180 degrees.
Any polygon with five or more sides having an angle less than 90 degrees between any two consecutive sides is unacceptable.
Examples of acceptable and non-acceptable polygons are shown below.


ACCEPTABLE
1
2
3
1
2
3
4
1
2
3
4
5
6
7
NON-ACCEPTABLE
1
2
3
4
5
6
7
8
9
10
11
12
13
14
1
2
3
4
5
6
7
8
9
Angle less than 90°

(Angles 8, 9, 10, 11
are greater than 180°)

The program gives users the flexibility for retrieving information for counties within a state or for counties within several states. Consequently, interstate AQCR's can be covered.

PROGRAMMING REQUIREMENTS - The following considerations apply only to the algorithm for polygon retrieval.

1. The number of points on the polygon is read followed by the coordinates of each point. The coordinates are stored in an array.

2. The coordinates of the polygon are compared to determine the minimum and maximum longitude (X) and latitude (Y) coordinates for the polygon.

3. The minimum and maximum X and Y become the end points for the sides of an imaginary rectangle enclosing the polygon.

4. Transform the latitude/longitude coordinates of the polygon to a rectangular coordinate system in kilometers. For simplicity, place the origin of the rectangular coordinate system at a point corresponding to the maximum longitude and minimum latitude of the n points of the polygon, i.e. start in the lower left corner of the imaginary rectangle defined in Step 3.
   Look up the length in kilometers of 1° of longitude for the latitude that lies half way between the minimum and maximum latitudes of the n points of the polygon. Table A.2.1 shows the lengths of 1° of longitude at various latitudes. The length of 1° of latitude is approximately 111.19 kilometers.

5. Calculate and store the slopes and intercepts of the n sides of the polygon. If the absolute value $(X_{i+1}) - X_i = 0$, then b_i = infinity. Test for $(X_{i+1}) - X_i = 0$ before calculating b_i, the slope of a line, in order to avoid dividing by zero

Table A.2.1 LENGTH OF A DEGREE OF LONGITUDE AT SPECIFIED LATITUDES

Lat°	Degree of Longitude (km)	Lat°	Degree of Longitude (km)
0	111.388	27	99.317
1	111.372	28	98.425
2	111.320	29	97.501
3	111.237	30	96.548
4	111.119	31	95.564
5	110.968	32	94.551
6	110.781	33	93.512
7	110.564	34	92.443
8	110.312	35	91.345
9	110.026	36	90.221
10	109.709	37	89.068
11	109.356	38	87.889
12	108.969	39	86.681
13	108.551	40	85.448
14	108.101	41	84.188
15	107.617	42	82.903
16	107.101	43	81.593
17	106.552	44	80.258
18	105.971	45	78.897
19	105.357	46	77.514
20	104.712	47	76.105
21	104.035	48	74.673
22	103.327	49	73.219
23	102.586	50	71.742
24	101.816	51	70.243
25	101.013	52	68.721
26	100.180	53	67.180

Table A.2.1 (continued) LENGTH OF A DEGREE OF LONGITUDE AT SPECIFIED LATITUDES

Lat°	Degree of longitude (km)	Lat°	Degree of longitude (km)
54	65.618	81	17.483
55	64.035	82	15.554
56	62.433	83	13.620
57	60.812	84	11.683
58	59.171	85	9.741
59	57.514	86	7.797
60	55.836	87	5.849
61	54.143	88	3.900
62	52.432	89	1.950
63	50.705		
64	48.963		
65	47.206		
66	45.435		
67	43.649		
68	41.849		
69	40.035		
70	38.211		
71	36.375		
72	34.527		
73	32.668		
74	30.800		
75	28.921		
76	27.034		
77	25.139		
78	23.234		
79	21.324		
80	19.406		

in the formula to calculate bi. If the value is zero, this indicates that the side of the polygon is perpendicular to the base. In testing for whatever point lies within the interval of the end points of a polygon side, we are actually testing to see if the longitude of the point lies within the interval defined by the maximum and minimum latitudes for the side.

6. Read the coordinates of a point source from NEDS or for an active monitoring site from the SAROAD Site Identification File.

7. If the coordinates of the point do not fall within the imaginary rectangle defined in Step 3 above, read the coordinates of the next NEDS point source or SAROAD monitoring site.

8. If the coordinates of the point fall within the imaginary rectangle, it is necessary to test to see if the point is within the defined polygon.

9. Transform the coordinates of the point in question to kilometers in the same manner as was done for the polygon points.

10. Read the coordinates of points i and i + 1 on the polygon.

11. If the X coordinate of the NEDS point source or SAROAD monitoring site, whichever is applicable, falls within the interval X_i to X_{i+1} on the polygon, it is necessary to determine the coordinate of the location on the side n of the polygon where an ordinate constructed at point X, the NEDS

point source or SAROAD monitoring site location, on the horizontal axis would intersect the polygon side. This point is defined as: Y(up) = $a_i + b_i X$.

12. Continue to read the coordinates of the end points of the remaining sides of the polygon to find the second side of the polygon where the ordinate at point X intersects the polygon. Define this point as: Y(low) = $a_j + b_j X$.

13. If $Y(\text{low}) \leq Y \leq Y(\text{up})$, then the coordinates of the point in question lie within the defined polygon.

Special formulae used are:

1. $|X_{i+1} - X_i| = 0$ tests the absolute value of the difference between two polygon points to determine if a slope and intercept can be calculated. This is to handle the special case in which a perfect vertical line has been defined. It is a necessary test to avoid dividing by zero in the next set of formulae.

2. $b_i = \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i}$ calculates the slope of each side of the polygon.

3. $a_i = Y_i - b_i X_i$ calculates the intercept for each side of the polygon along the Y axis of the starting point, i.e. the point with the largest longitude coordinate.

The accompanying diagrams illustrate the process. The notations are:

X_i = The largest longitude (X) coordinate of the ith side of the polygon.

X_{i+1} = the smallest longitude (X) coordinate of the ith side of the polygon.

Y up = the highest Y point at which a vertical line through a point within the polygon intersects the ith side of the polygon.

Y low = the lowest Y point at which a vertical line through a point within the polygon intersects the ith side of the polygon.

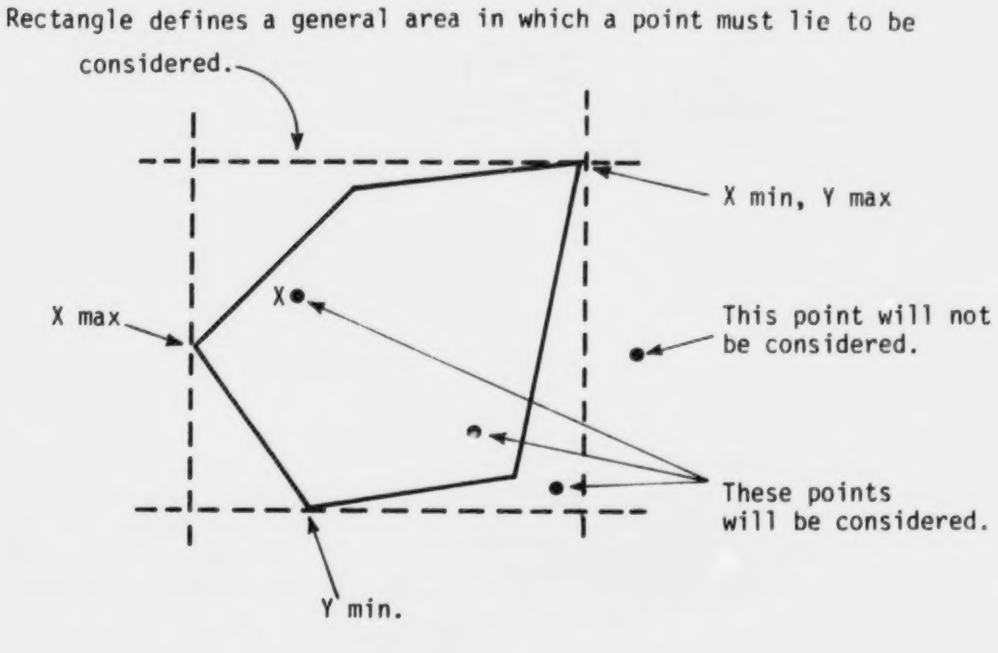
X = the point being tested from the USER file or from the SAROAD Site Identification file.

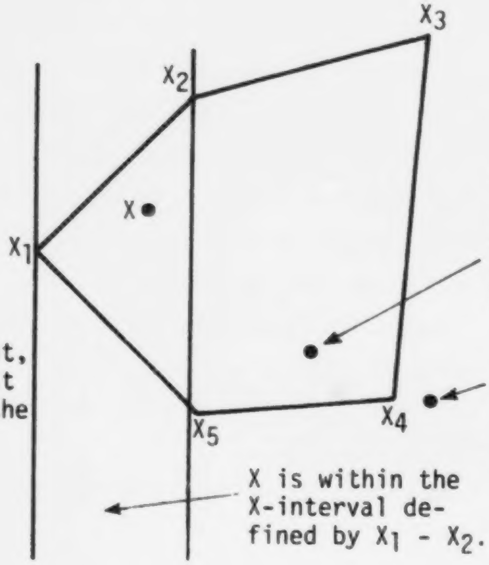
X min. = smallest longitude (X) coordinate value for the polygon.

X max. = largest longitude (X) coordinate value for the polygon.

Y min. = smallest latitude (Y) coordinate value for the polygon.

Y max. = largest latitude (Y) coordinate value for the polygon.


Rectangle defines a general area in which a point must lie to be considered.
X min, Y max
X
X max
This point will not be considered.
These points will be considered.
Y min.

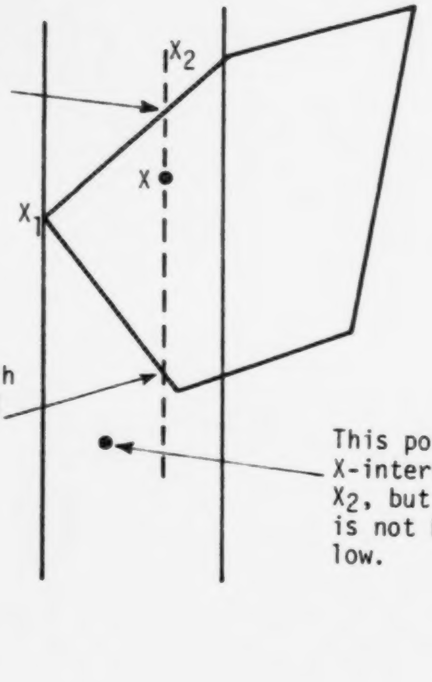


X_1 is the starting point, since it is the smallest X coordinate value on the polygon.

This point is within an X-interval that is defined by X_2 - X_3 and also within the interval defined by X_5 - X_4.

This point is not within a polygon X-interval.

X is within the X-interval defined by X_1 - X_2.



Y up = Y value at which vertical line through X intersects line.

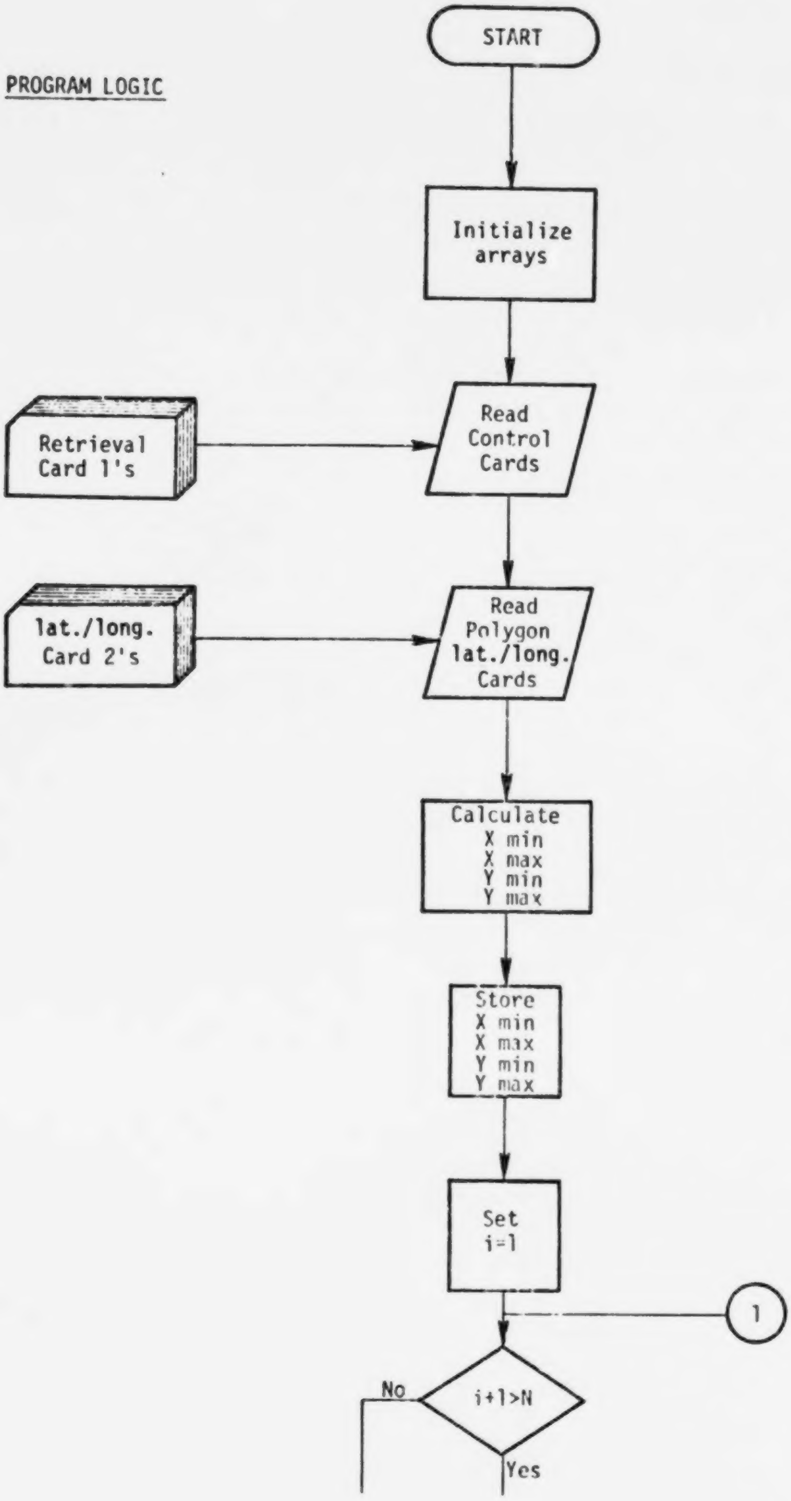
Y low = Y value at which vertical line through X intersects line.

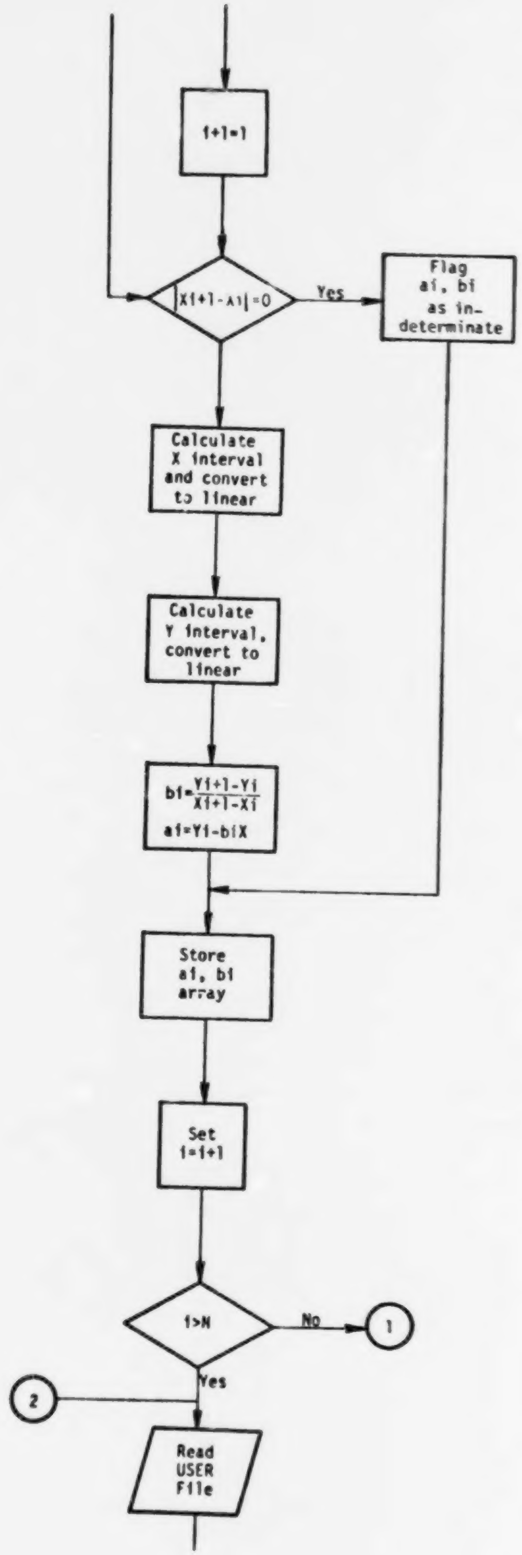
This point lies within the X-interval defined by X_1 - X_2, but its Y coordinate is not between Y up and Y low.

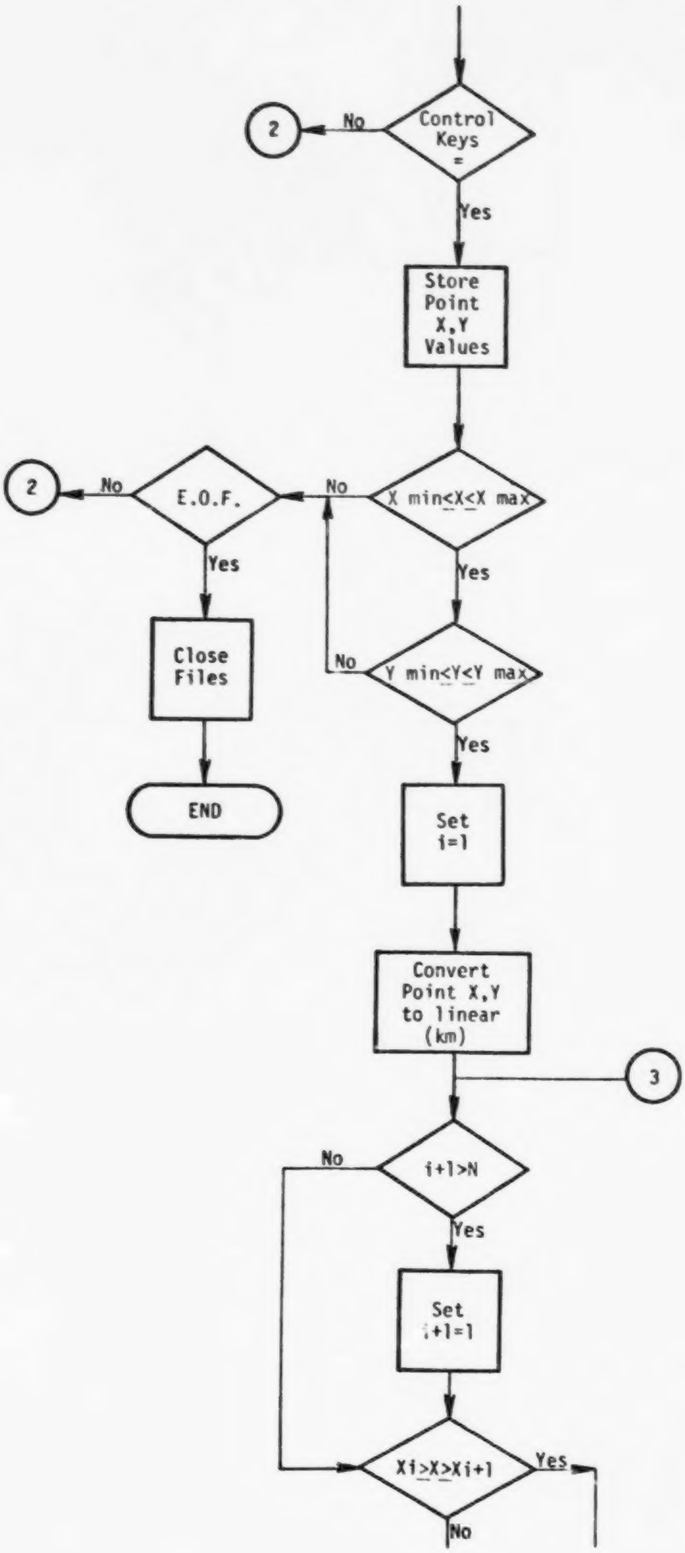
LEVEL OF EFFORT and ANTICIPATED COST -

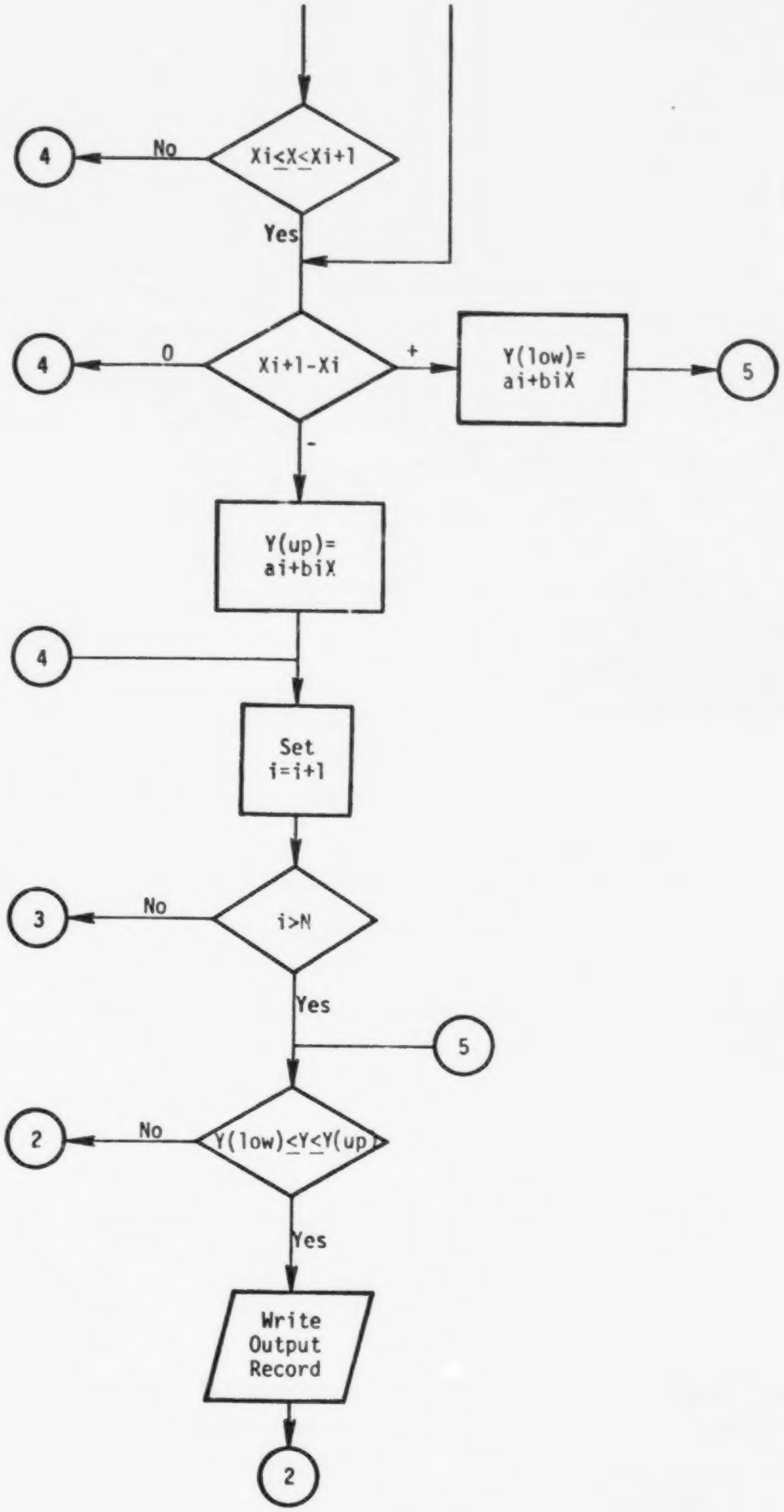
1.	Preliminary study and problem definition = 160 hrs @ $20/hr.	=	$ 3200.00
2.	Coding = 460 hours @ $20/hr. plus 64 hours clerical support @ $5/hr.	=	9520.00
3.	Testing and debugging = 480 hours @ $20/hr.	=	9600.00
4.	Documentation = 70 hours @ $20/hr. plus 40 hours clerical support @ $5/hr.	=	1600.00
			23,920.00
5.	Computer time is		
	12 compile/test runs @ 20 minutes/run = 240 min = 4 SUP @ $368/SUP	=	1475.00
			$25,395.00

PROGRAM LOGIC


START
Initialize arrays
Retrieval Card 1's
Read Control Cards
lat./long. Card 2's
Read Polygon lat./long. Cards
Calculate X min X max Y min Y max
Store X min X max Y min Y max
Set i=1
1
i+1>N
No
Yes


i+1=1
|Xi+1-Xi|=0
Yes
Flag ai, bi as indeterminate
Calculate X interval and convert to linear
Calculate Y interval, convert to linear
bi=(Yi+1-Yi)/(Xi+1-Xi)
ai=Yi-biX
Store ai, bi array
Set i=i+1
i>N
No
1
Yes
2
Read USER File


Control Keys =
No
2
Yes
Store Point X,Y Values
X min≤X≤X max
No
E.O.F.
No
2
Yes
Close Files
END
Yes
Y min≤Y≤Y max
No
Yes
Set i=1
Convert Point X,Y to linear (km)
3
i+1>N
No
Yes
Set i+1=1
Xi≥X≥Xi+1
Yes
No


4
No
Xi≤X≤Xi+1
Yes
4
0
Xi+1-Xi
+
Y(low)=
ai+biX
5
-
Y(up)=
ai+biX
4
Set
i=i+1
3
No
i>N
Yes
5
2
No
Y(low)≤Y≤Y(up)
Yes
Write
Output
Record
2

A.3.

APPLICATION - Parametric data retrieval.

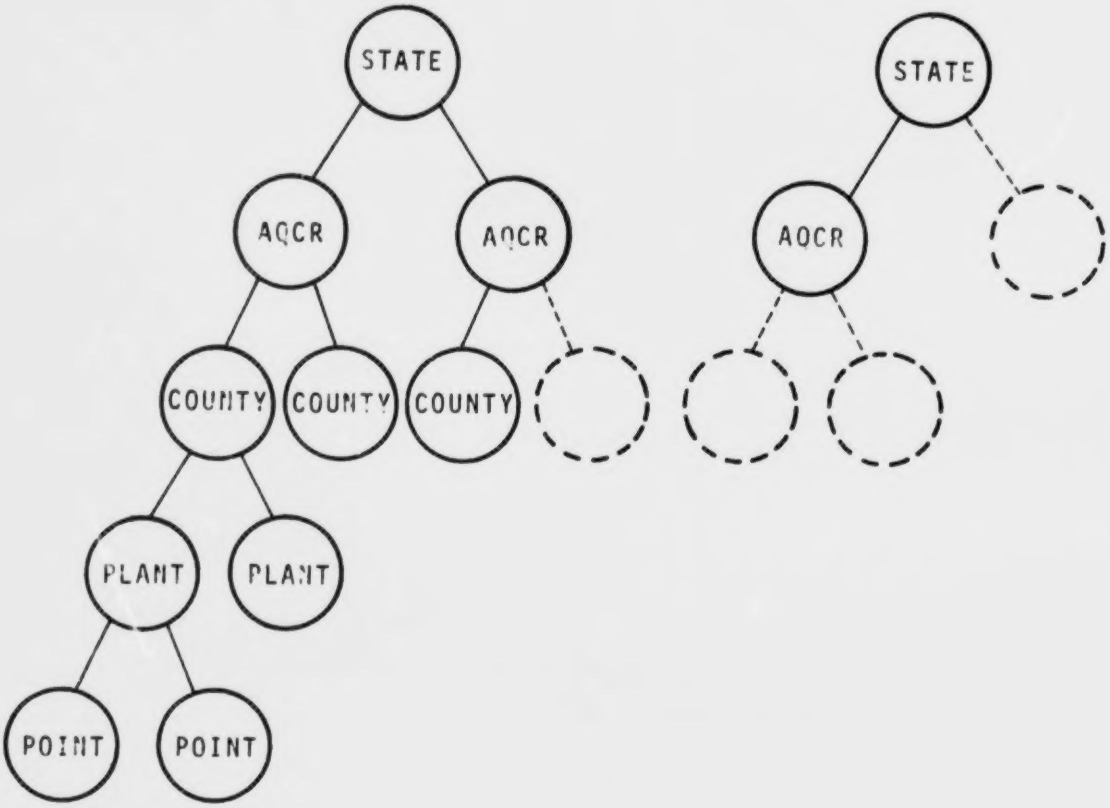
ABSTRACT - The purpose of this program is to allow NEDS users to specify a value range for one or more NEDS point source parameters and to retrieve data only for those sources having values within the range. The user could specify his own report design, or he could retrieve applicable plant/point source numbers with which to access existing retrieval programs such as NE214, the Point Source Listing program. This is an application for SYSTEM 2000 or an equivalent file management system. The specifications included here are for SYSTEM 2000. The application has two phases. Phase 1 is a program written in Procedural - Language - Feature - COBOL to establish a subfile of the NADB* NEDS - USER file. The Phase 1 programming should be completed by NADB to avoid duplication of effort within other EPA offices. Phase 2 is a SYSTEM 2000 program written to retrieve data elements in IMMEDIATE ACCESS mode from the data base created in Phase 1. The user will supply the commands necessary to load the data base and to interrogate the elements. The rationale for this application is that it can improve the applicability of NEDS for special studies by eliminating the time needed for manual review of reports. The total report print time could be reduced, thereby improving report turnaround time.

INPUT DESCRIPTION - Input to Phase 1 is the NADB* NEDS -USER file, Input to Phase 2 is the SYSTEM 2000 defined file created in Phase 1.

OUTPUT DESCRIPTION - Output from Phase 1 is a data base with the following description.

The heirarchical levels are:

Level 0 = State
Level 1 = AQCR
Level 2 = County
Level 3 = Plant
Level 4 = Point


STATE
STATE
AQCR
AQCR
AQCR
COUNTY
COUNTY
COUNTY
PLANT
PLANT
POINT
POINT

The KEY elements are:

State Number
AQCR Number
County Number
Plant Number
Point Number
SIC
Ownership Code
Stack Height
Boiler Design Capacity
Control Equipment Codes
Emissions Estimates
Estimation Methods
Compliance Schedule
SCC
Percent Sulfur
Operating Rate
Maximum Design Rate
Percent Ash

The data base definition is:

```
 1* STATE (KEY NAME XX):
 2* AQCR (RG):
    3* AQCR-NUMBER (KEY NAME XXX IN 2):
    4* COUNTY (RG IN 3):
       5* COUNTY-NUMBER (KEY NAME XXXX IN 4):
       6* PLANT (RG IN 5):
          7* PLANT-NUMBER (KEY INTEGER NUMBER XXXX IN 6):
          8* PLANT-NAME (NAME X(40) IN 6):
          9* OWNERSHIP (KEY NAME X IN 6):
         10* POINT (RG IN 6):
```

```
11* POINT-NUMBER (KEY INTEGER NUMBER 99 IN 10):
12* SIC (KEY INTEGER NUMBER 9(4) IN 10):
13* STACK-HEIGHT (KEY INTEGER NUMBER 9(4) IN 10):
14* STACK-DIAMETER (NON-KEY DECIMAL NUMBER 99.9 IN 10):
15* STACK-TEMP (NON-KEY INTEGER NUMBER 9(4) IN 10):
16* FLOW (NON-KEY INTEGER NUMBER 9(7) IN 10):
17* PLUME (NON-KEY INTEGER NUMBER 9(4) IN 10):
18* BOILER (KEY INTEGER NUMBER 9(4) IN 10):
19* PRIM-PART (KEY INTEGER NUMBER 9(3) IN 10):
20* SEC-PART (KEY INTEGER NUMBER 9(3) IN 10):
21* PRIM-SOX (KEY INTEGER NUMBER 9(3) IN 10):
22* SEC-SOX (KEY INTEGER NUMBER 9(3) IN 10):
23* PRIM-NOX (KEY INTEGER NUMBER 9(3) IN 10):
24* SEC-NOX (KEY INTEGER NUMBER 9(3) IN 10):
25* PRIM-HC (KEY INTEGER NUMBER 9(3) IN 10):
26* SEC-HC (KEY INTEGER NUMBER 9(3) IN 10):
27* PRIM-CO (KEY INTEGER NUMBER 9(3) IN 10):
28* SEC-CO (KEY INTEGER NUMBER 9(3) IN 10):
29* PART-EFFIC (KEY DECIMAL NUMBER 99.9 IN 10):
30* SOX-EFFIC (KEY DECIMAL NUMBER 99.9 IN 10):
31* NOX-EFFIC (KEY DECIMAL NUMBER 99.9 IN 10):
32* HC-EFFIC (KEY DECIMAL NUMBER 99.9 IN 10):
33* CO-EFFIC (KEY DECIMAL NUMBER 99.9 IN 10):
34* PERCENT-WINTER (NON-KEY INTEGER NUMBER 99 IN 10):
35* PERCENT-SPRING (NON-KEY INTEGER NUMBER 99 IN 10):
36* PERCENT-SUMMER (NON-KEY INTEGER NUMBER 99 IN 10):
37* PERCENT-FALL (NON-KEY INTEGER NUMBER 99 IN 10);
38* HOURS-DAY (NON-KEY INTEGER NUMBER 99 IN 10):
39* DAYS (NON-KEY INTEGER NUMBER 9 IN 10):
40* WEEKS (NON-KEY INTEGER NUMBER 99 IN 10):
41* PART-EST (KEY INTEGER NUMBER 9(7) IN 10):
42* SOX-EST (KEY INTEGER NUMBER 9(7) IN 10):
43* NOX-EST (KEY INTEGER NUMBER 9(7) IN 10):
44* HC-EST (KEY INTEGER NUMBER 9(7) IN 10):
45* CO-EST (KEY INTEGER NUMBER 9(7) IN 10):
```

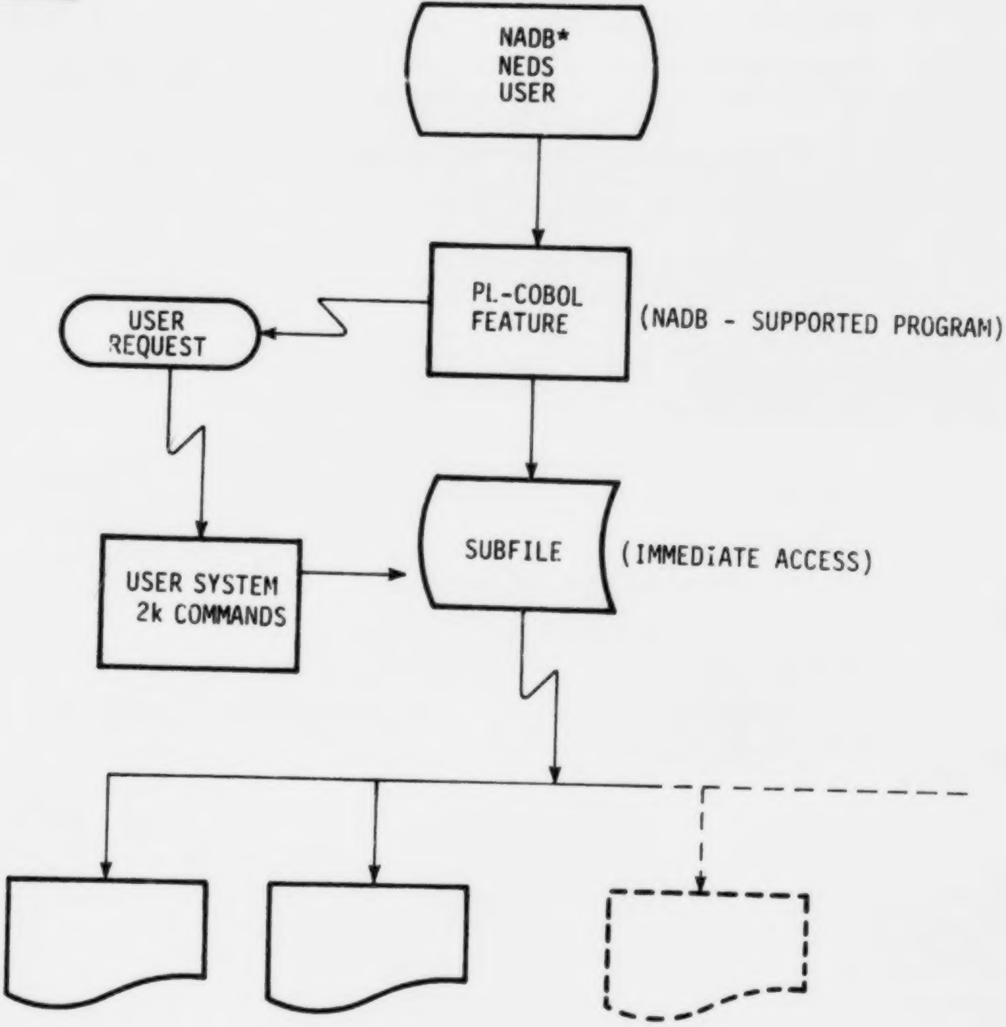
```
46* PART-METH (KEY INTEGER NUMBER 9 IN 10):
47* SOX-METH (KEY INTEGER NUMBER 9 IN 10):
48* NOX-METH (KEY INTEGER NUMBER 9 IN 10):
49* HC-METH (KEY INTEGER NUMBER 9 IN 10):
50* CO-METH (KEY INTEGER NUMBER 9 IN 10):
51* SPACE-HEAT (NON-KEY DECIMAL NUMBER 99.9 IN 10):
52* PART-ALLOWABLE (KEY INTEGER NUMBER 9(7) IN 10):
53* SOX-ALLOWABLE (KEY INTEGER NUMBER 9(7) IN 10):
54* NOX-ALLOWABLE (KEY INTEGER NUMBER 9(7) IN 10):
55* HC-ALLOWABLE (KEY INTEGER NUMBER 9(7) IN 10);
56* CO-ALLOWABLE (KEY INTEGER NUMBER 9(7) IN 10):
57* SOURCE (RG IN 10):
58* SCC-I (KEY INTEGER NUMBER 9 IN 57):
59* SCC-II (KEY INTEGER NUMBER 999 IN 57):
60* SCC-III (KEY INTEGER NUMBER 9(6) IN 57):
61* SCC-IV (KEY INTEGER NUMBER 9(8) IN 57):
62* OPERATING-RATE (KEY INTEGER NUMBER 9(7) IN 57):
63* MAX-DESIGN (KEY DECIMAL NUMBER 9999.999 IN 57):
64* SULFUR (KEY DECIMAL NUMBER 9.99 IN 57):
65* ASH (KEY DECIMAL NUMBER 99.9 IN 57):
```

Output from Phase 2 is defined by the user.

OPERATION - Input is user supplied commands to load NADB* NEDS USER file or portions of it into the data base. The user then provides a set of commands to interrogate the data base and retrieve reports for which he specifies the format.

PROGRAMMING REQUIREMENTS - The IMMEDIATE access mode is specified because of the features available. The MIN, MAX functions of the WHERE clauses or the Ternary Operators EQ.NE. SPANS are necessary for maximum benefits to be derived by this application.

SYSTEM FLOW


NADB*
NEDS
USER
PL-COBOL
FEATURE
(NADB - SUPPORTED PROGRAM)
USER
REQUEST
SUBFILE
(IMMEDIATE ACCESS)
USER SYSTEM
2k COMMANDS

USER - DEFINED REPORTS

LEVEL OF EFFORT and ANTICIPATED COST -

1.	Preliminary study and problem definition = 120 hours @ $20/hr.	=	$2,400.00
2.	Structuring and coding the data base = 40 hours @ $20/hr. plus 16 hours clerical support @ $5/hr.	=	880.00
3.	Testing and debugging data base loading = 80 hours @ $20/hr.	=	1,600.00
4.	Coding retrieval logic = 40 hours @ $20/hr. plus 16 hours clerical support $5/hr.	=	880.00
5.	Testing and debugging retrieval logic = 40 hours @ $20/hr.	=	800.00
6.	Documentation = 24 hours @ $20/hr plus 16 hours clerical support @ $5/hr.	=	560.00
			$7,120.00
7.	Computer time is:		
	Data base loading debug = 10 tests @ 15 min/test = 150 min = 2.5 SUP @ $368/SUP	=	$ 920.00
	Retrieval testing = 5 tests @ 10 min/test = 50 min = .8 SUP @ $368/SUP	=	$ 295.00
			$1,215.00
		TOTAL	$8,335.00

A.4

APPLICATION - Analysis of the effects of potential changes to NEDS data.

ABSTRACT - The purpose of this program is to allow users to substitute new values for parameters in an existing point source record and to analyze the effects on emissions. For example, the user may want to see the effect on emissions if all sources of a specific type were required to burn fuel of a specified sulfur content. This application can be a file-management system application or a COBOL application. A SYSTEM 2000 approach involves several job steps, whereby the COBAL application is less cumbersome and can work using control cards if the job application requirements can be properly defined prior to programming. Flow diagrams are presented here for both concepts.

PROGRAM SPECIFICATIONS - The following description applies to a single COBOL program for this application.

1. Selection criteria are:

   State
   County
   AQCR
   Plant
   SCC

2. Parameters that can be specified to be changed in each record are:

   Estimated Control Efficiency
   Operating Rate
   Sulfur Content
   Any combination of the above

3. After the specified parameter has been changed in the record, the emission factor will be applied, and the emissions will be recalculated.

4. Output will show:

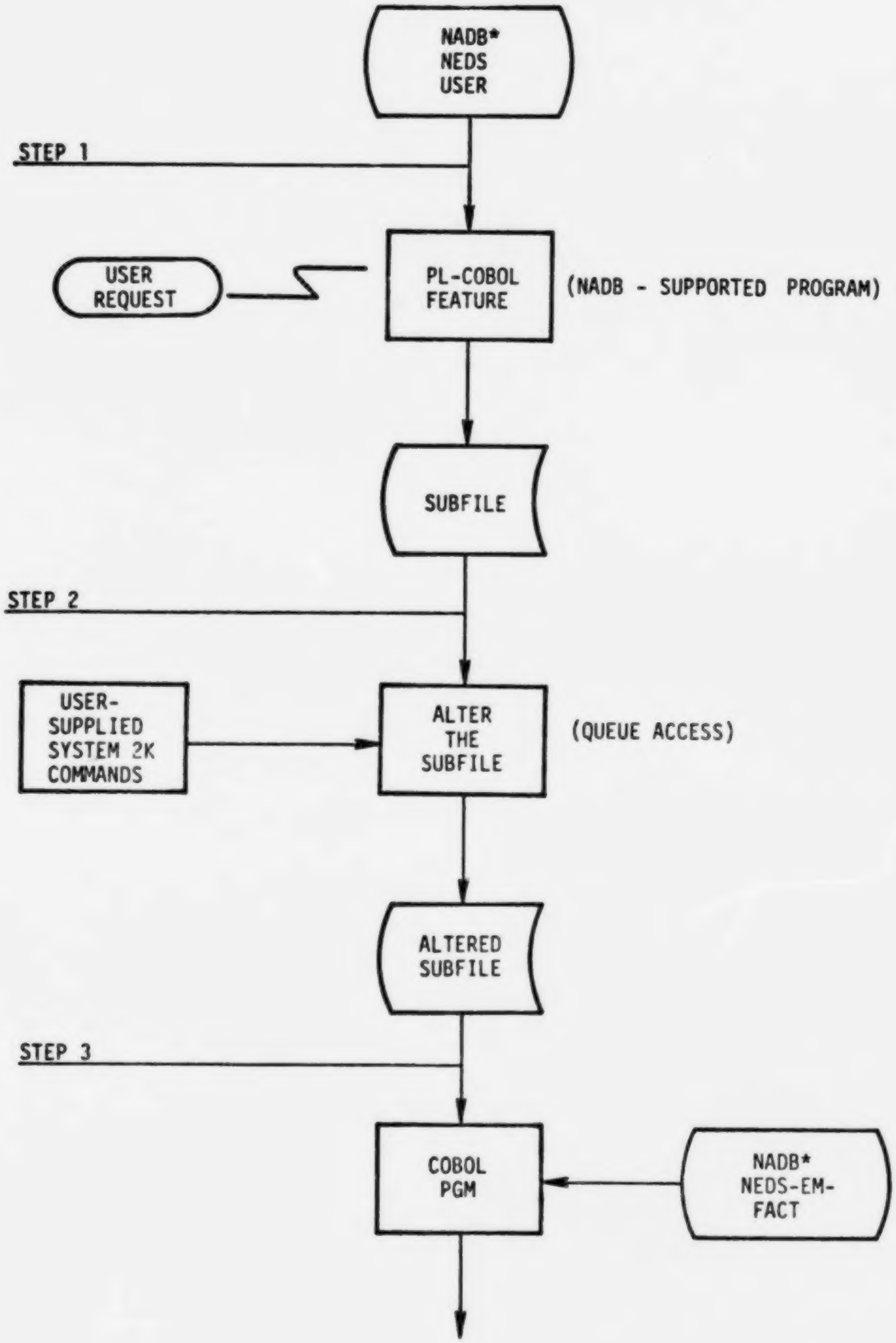
   Header with control card data
   State, County, AQCR number
   Plant I.D.
   Plant name and address
   Emissions by point

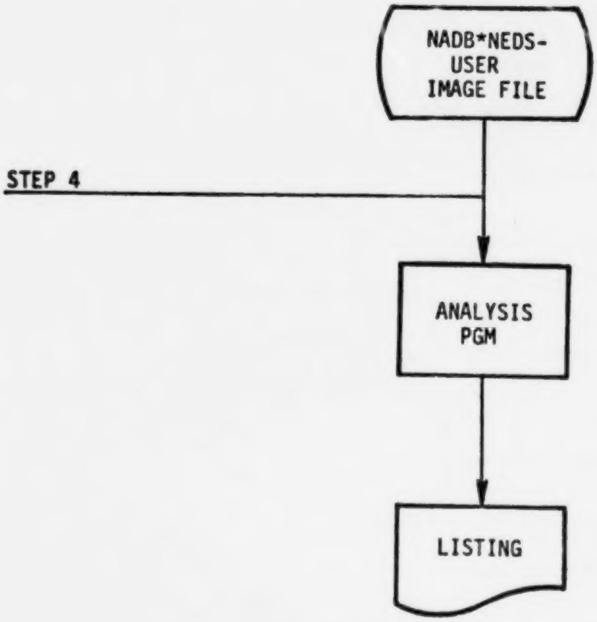
5. Program can be altered to punch AQDM or CDM - format input cards if desired.

LEVEL OF EFFORT
and
ANTICIPATED COST -

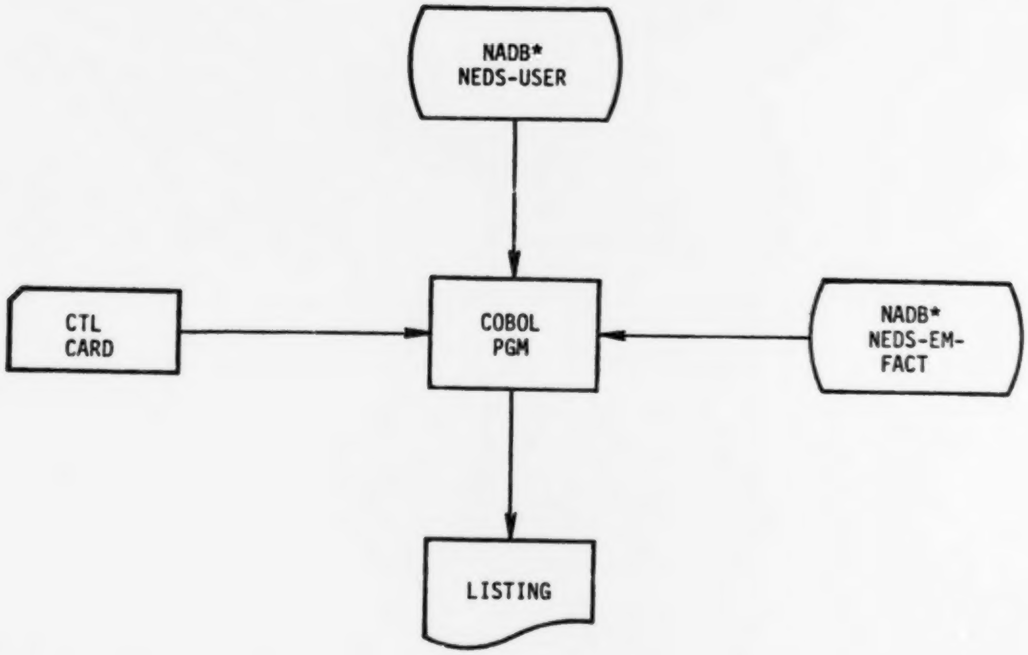
1.	Preliminary study and problem definition = 160 hours @ $20/hr.	=	$3,200.00
2.	Program coding and testing 160 hours @ $20/hr. plus 40 clerical hours @ $5/hr.	=	3,400.00
3.	Documentation = 40 hours @ $20/hr. plus 20 hours @ $5/hr.	=	900.00
			$7,500.00
4.	Computer time is:		
	8 compile/test runs @ 20 min/run = 160 min = 2.75 SUP @ $368/SUP	=	$1,015.00
			$8,515.00

SYSTEM FLOW (USING SYSTEM 2000)

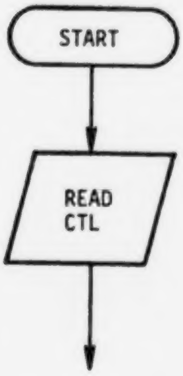

NADB*
NEDS
USER
STEP 1
USER
REQUEST
PL-COBOL
FEATURE
(NADB - SUPPORTED PROGRAM)
SUBFILE
STEP 2
USER-
SUPPLIED
SYSTEM 2K
COMMANDS
ALTER
THE
SUBFILE
(QUEUE ACCESS)
ALTERED
SUBFILE
STEP 3
COBOL
PGM
NADB*
NEDS-EM-
FACT

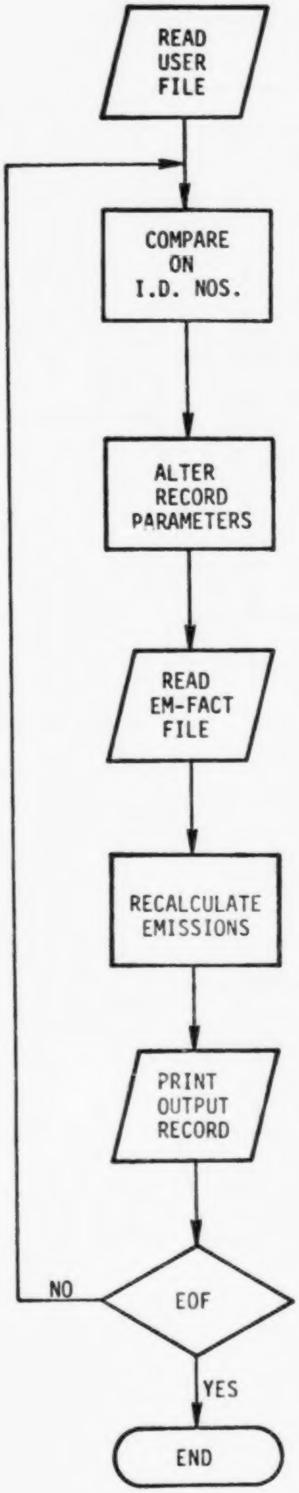

NADB*NEDS-
USER
IMAGE FILE
STEP 4
ANALYSIS
PGM
LISTING

SYSTEM FLOW (USING COBOL)


NADB*
NEDS-USER
CTL
CARD
COBOL
PGM
NADB*
NEDS-EM-
FACT
LISTING

PROGRAM FLOW (COBOL)


START
READ
CTL


READ USER FILE
COMPARE ON I.D. NOS.
ALTER RECORD PARAMETERS
READ EM-FACT FILE
RECALCULATE EMISSIONS
PRINT OUTPUT RECORD
EOF
NO
YES
END

A.5

APPLICATION - Basic accounting capability.

ABSTRACT - The purpose of this application is to allow NEDS users to determine the number of sources or facilities in the file that conform to some defined criteria, e.g. the number of coal-fired boilers in the file or the number of sources with boilers burning multiple fuels. The application is for SYSTEM 2000 or an equivalent file management system. The specifications included here are for SYSTEM 2000. The application has two phases. Phase 1 is a program written in Procedural Language - Feature - COBOL to establish a subfile of the NADB*NEDS-USER file. The Phase 1 programming should be completed by NADB to avoid duplication of effort within other EPA offices. Phase 2 is a SYSTEM 2000 program written to retrieve data elements in IMMEDIATE ACCESS mode from the data base created in Phase 1. The user will supply the commands necessary to load the data base and to interrogate the file. The rationale for this application is that manpower should be saved in scanning reports.

INPUT DESCRIPTION - Input to Phase 1 is the NADB*NEDS-USER file. Input to Phase 2 is the SYSTEM 2000 defined file created in Phase 1.

OUTPUT DESCRIPTION - Output from Phase 1 is a data base with the following description. The hierarchical levels are:

Level 0 = State
Level 1 = AQCR
Level 2 = County
Level 3 = Plant
Level 4 = Point

The KEY elements are:

- State Numbers
- AQCR Numbers
- County Numbers
- SIC
- Ownership Code
- Stack Height
- Boiler Design Capacity
- Control Equipment Codes
- Control Efficiencies
- Emissions Estimates
- Estimation Methods
- Compliance Schedule
- Regulations
- SCC
- Percent Sulfur
- Percent Ash
- Operating Rate
- Maximum Design Rate

The data base definition is the same as the one described for parametric data retrieval in Appendix A.3. The difference is in the addition of The Regulations as KEY elements in this data base. The most economical approach to this application and the one in Appendix A.3 is to have one data base defined that will fulfill the requirements for both applications.

OPERATION - Input is user supplied commands to load NADB*NEDS-USER file or portions of it into the data base. The user then provides a set of commands to interrogate the data base.

PROGRAMMING REQUIREMENTS - The TALLY and SUM features available from SYSTEM 2000 in the IMMEDIATE ACCESS mode are the primary vehicles for this application.

LEVEL OF EFFORT
and
ANTICIPATED COST - (Assumes parametric data retrieval base is built)

1.	Preliminary study and problem definition = 80 hours @ $20/hr.	=	$1,600.00
2.	Testing and debugging data base loading = 40 hours @ $20/hr.	=	800.00
3.	Coding retrieval logic = 60 hours @ $20/hr. plus 8 hours clerical support @ $5/hr.	=	1,240.00
4.	Testing and debugging = 40 hours @ $20/hr.	=	800.00
5.	Documentation = 16 hours @ $20/hr. plus 16 hours clerical support @ $5/hr.	=	400.00
			$4,840.00
6.	Computer time is:		
	Testing retrieval logic = 7 runs @ 25 min/run = 3 SUP @ $368/SUP	=	$1,105.00
			$5,945.00

A.6

APPLICATION - NEDS-CDS cross-reference.

ABSTRACT - The purpose of this program is to create and maintain a cross-reference file for NEDS and CDS plant and point source numbers. The capability to print the file in either NEDS or CDS sequence is incorporated. The successful use of this program requires that both NADB and DSSE update the cross-reference file at regular intervals. The rationale for the program is two-fold. First, NEDS might become more useful to DSSE personnel if correspondence between NEDS and CDS can be shown in the NEDS files. Second, the cross-reference capability is a potential source for indicating to both systems when new source data might be available. The program is in COBOL. A feasibility study is required to determine further requirements beyond the scope of this description.

INPUT DESCRIPTION - Input to this program comes from both NADB (NEDS) and DSSE (CDS). Input from NEDS is Cards 1 and 2 submitted with NEDS updates. Card 1 will be used to indicate a plant that has been added to, or deleted from, NEDS. Card 2 will be used to indicate points that have been added to, or deleted from, NEDS.

Input from CDS is Cards 2 and 5, the cards that are used to input CDS source information and emission point information. The combination of these cards is necessary in order to cross-reference at both the plant and point level. The formats for Cards 2 and 5 follow:

Card 2

CC	Field Description	Picture
1-2	Region Code	99
3-4	State Code	99
5-8	SAROAD County Code	9(4)
9-13	CDS Source Code	9(5)
14-54	Filler	X(41)
55-58	NEDS Plant No. (if different from CDS Source Code)	9(4)
59-79	Filler	X(21)
80	Transaction Code	X

Card 5

CC	Field Description	Picture
1-2	Region Code	99
3-4	State Code	99
5-8	SAROAD County Code	9(4)
9-13	CDS Source Code	9(5)
14-16	CDS Point Code	9(3)
17-19	Filler	X(3)
20-27	SCC	9(8)
28-29	NEDS Point No.	99
30-79	Filler	X(50)
80	Transaction Code	X

OUTPUT DESCRIPTION - Output from this program can be:

1. An updated Cross-Reference File
2. File listings in NEDS or CDS sequence

1. Updated Cross-Reference File - Output from the update option is a new Cross-Reference File. The file is a tape file with sequential, variable length records in 2 formats

Plant Format

Position No.	Description
1-2	State No.
3-6	County No.
7-9	AQCR No.
10-13	NEDS Plant I.D.
14-18	CDS Source No.
19-30	Blank

Point Format

1-2	NEDS Point I.D.
3-5	CDS Point I.D.
.	.
.	.
.	.

Repeat the format for each NEDS Point I.D. within a NEDS Plant I.D.
Records are sorted by NEDS Point I.D. within NEDS Plant I.D.

When NEDS cards are used to update the cross-reference file, they are compared to the cross-reference file to determine if an existing NEDS-CDS cross-reference record exists. If the record does not exist, the NEDS Plant and Point I.D. is printed. If the record does exist, then the record is deleted if the NEDS transaction type is "D". If the NEDS delete card is Card 1, the entire plant is deleted. If the NEDS delete card is Card 2, the point is deleted.

Cross-Reference Listing - NEDS Sequence Format

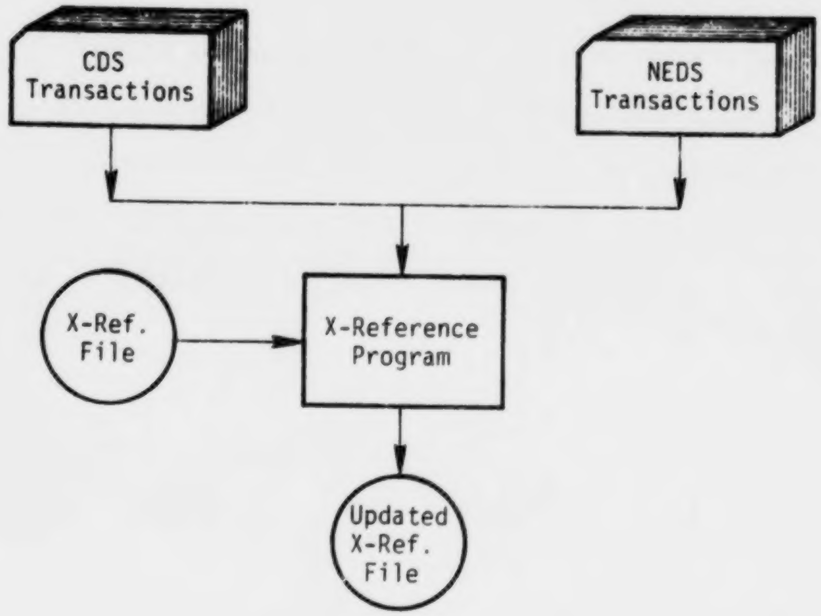
```
           NEDS-CDS CROSS-REFERENCE ------- UPDATES FROM XX XX

           NEDS IDENTIFICATION                CDS IDENTIFICATION

STATE    COUNTY    AQCR    PLANT    POINT    SOURCE    POINT

 XX       XXXX      XXX     XXXX     XX       XXXXX     XX
                                                        XX
                                                        XX

                                     XX                 XX
                                                        XX
                                                        XX

                            XXXX     XX       XXXXX     XX
                                                        XX
```

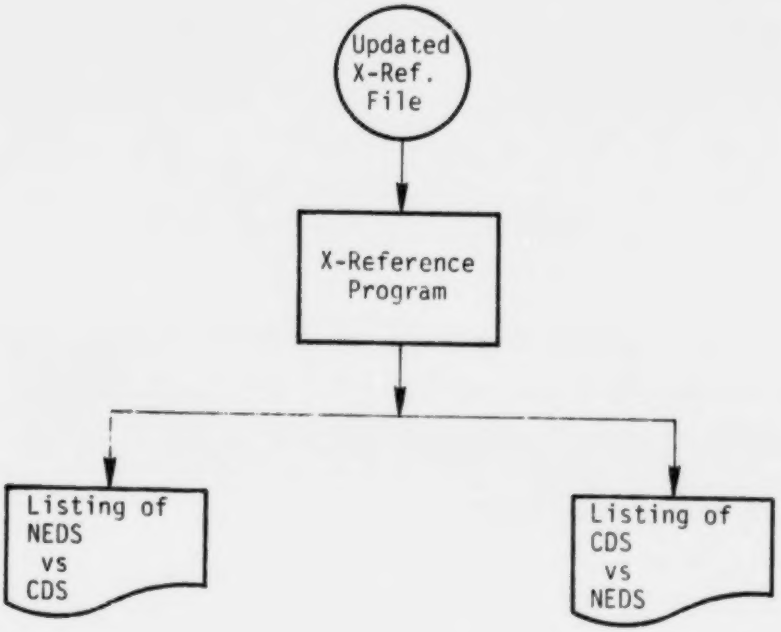
Cross-Reference Listing - CDS Sequence Format

```
           CDS-NEDS CROSS-REFERENCE ------- UPDATES FROM XX XX

           CDS IDENTIFICATION            NEDS IDENTIFICATION
STATE        COUNTY       SOURCE       POINT       PLANT       POINT

 XX           XXXX         XXXXX        XX          XXXX        XX
                                        XX
                                        XX
                                        XX

                           XXXXX        XX          XXXX        XX
                                        XX
```

<u>OPERATION</u> - Update transactions will be applied against the Cross-Reference File at regular intervals. A control card is used to define the run options. The options are:

1. Cross-check from NEDS
2. Update from CDS
3. Update and print from CDS
4. Print in NEDS sequence
5. Print in CDS sequence
6. Select up to eight states for the above functions
7. Perform the above functions for all states.


CDS Transactions
NEDS Transactions
X-Ref. File
X-Reference Program
Updated X-Ref. File


Updated X-Ref. File
X-Reference Program
Listing of NEDS vs CDS
Listing of CDS vs NEDS

The control card format is:

CC	Valid Codes	Function
1-3	CTL	Identifies control card.
4	Update option	
	1	Update from NEDS.
	2	Update from CDS.
5	Print option	
	1	Print in NEDS sequence.
	2	Print in CDS sequence.
6-9	Month, Year	Identifies month and year of update run. On updates, the month and year are added to the end of any affected plant record. If the update option (CC4) is blank, the month and year entered in CC 6-9 will cause only those records updated after and including that date to be printed.
10-11	State No.	Identifies the state or states to be printed.
	Default=blanks	All states will be printed.
12-25	Repeat state numbers until up to eight states have been entered. Designation of states applies only to the print options.	

Input transaction cards must be sorted as follows:

NEDS TRANSACTIONS

- State — MAJOR
  - County
    - AQCR — MINOR
      - Plant
        - Point

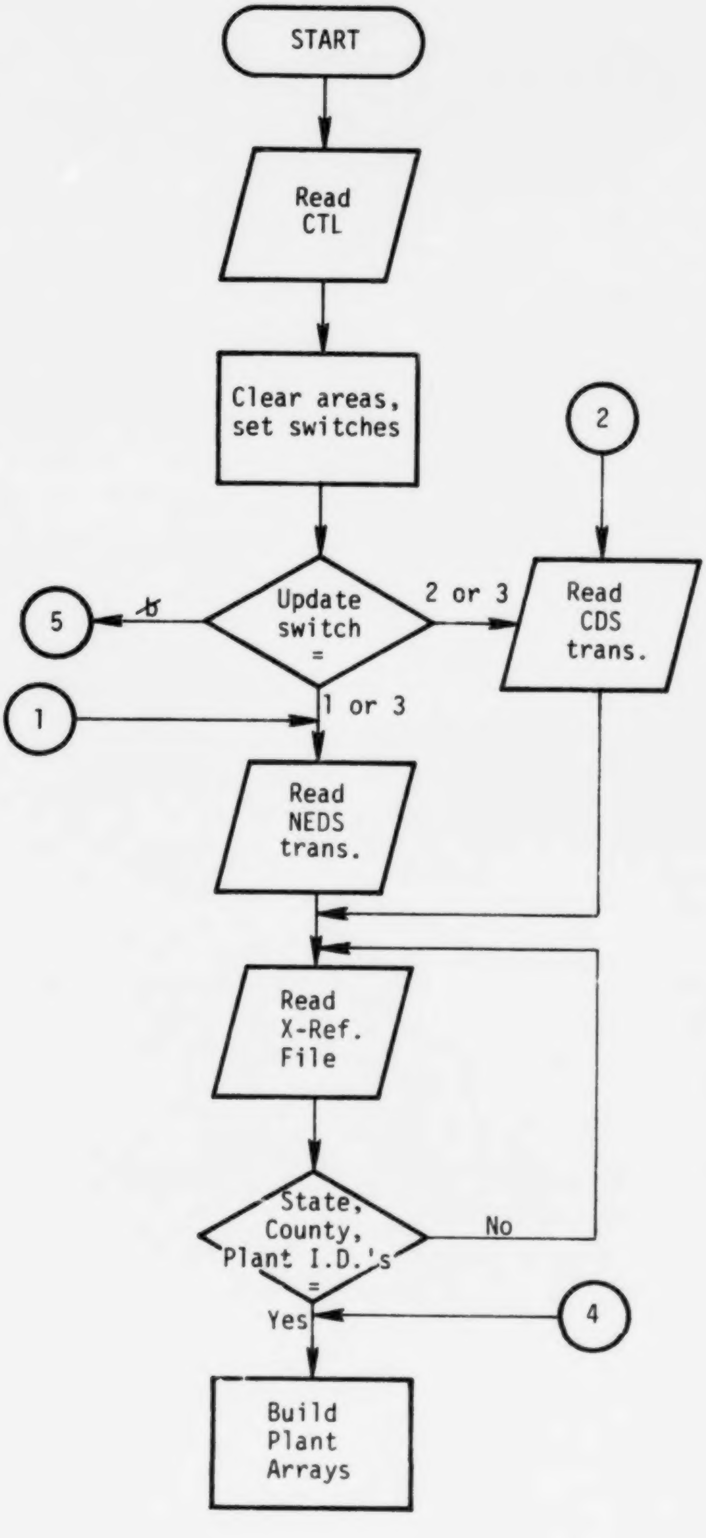
CDS TRANSACTIONS

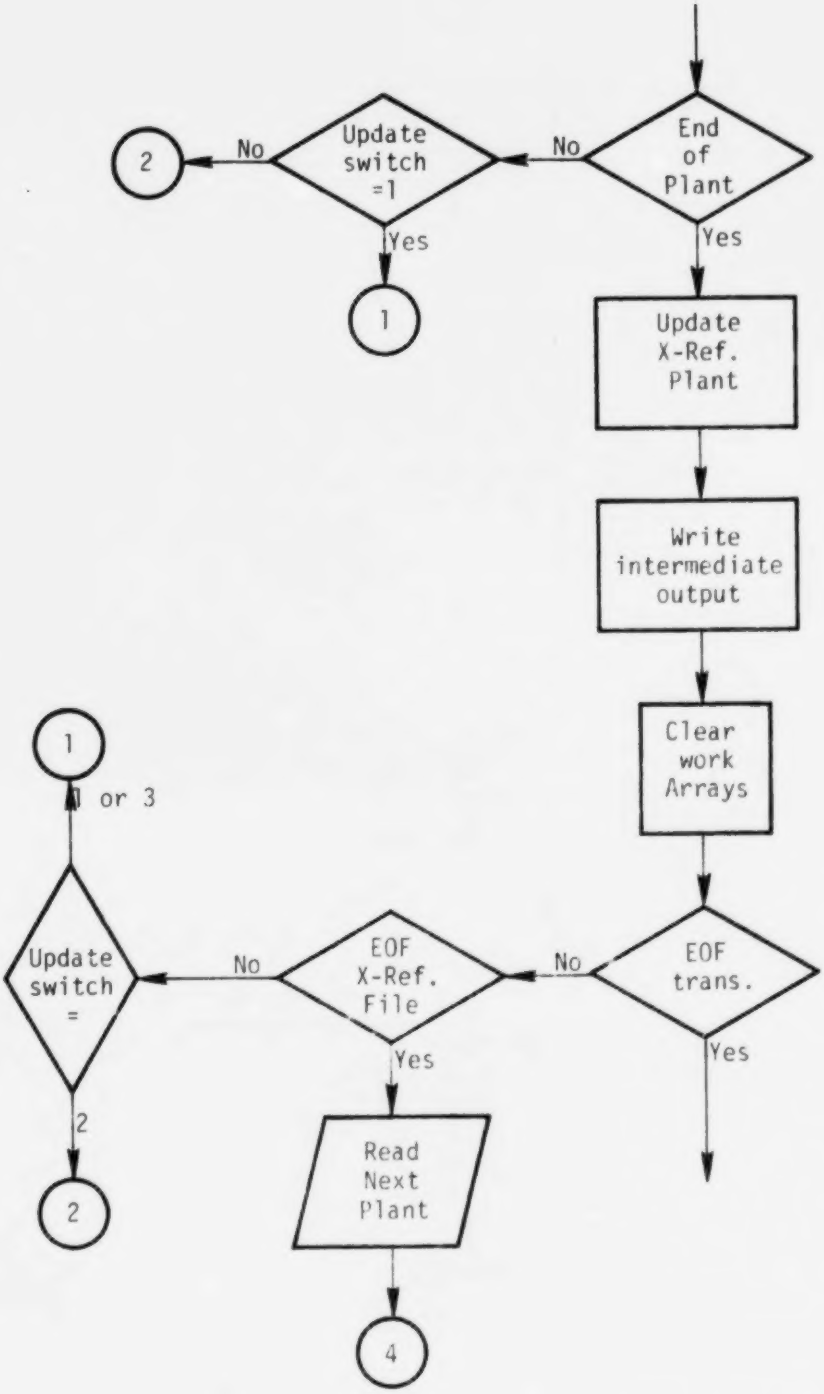
- State — MAJOR
  - County
    - Source — MINOR
      - Emission Point

LEVEL OF EFFORT and ANTICIPATED COST -

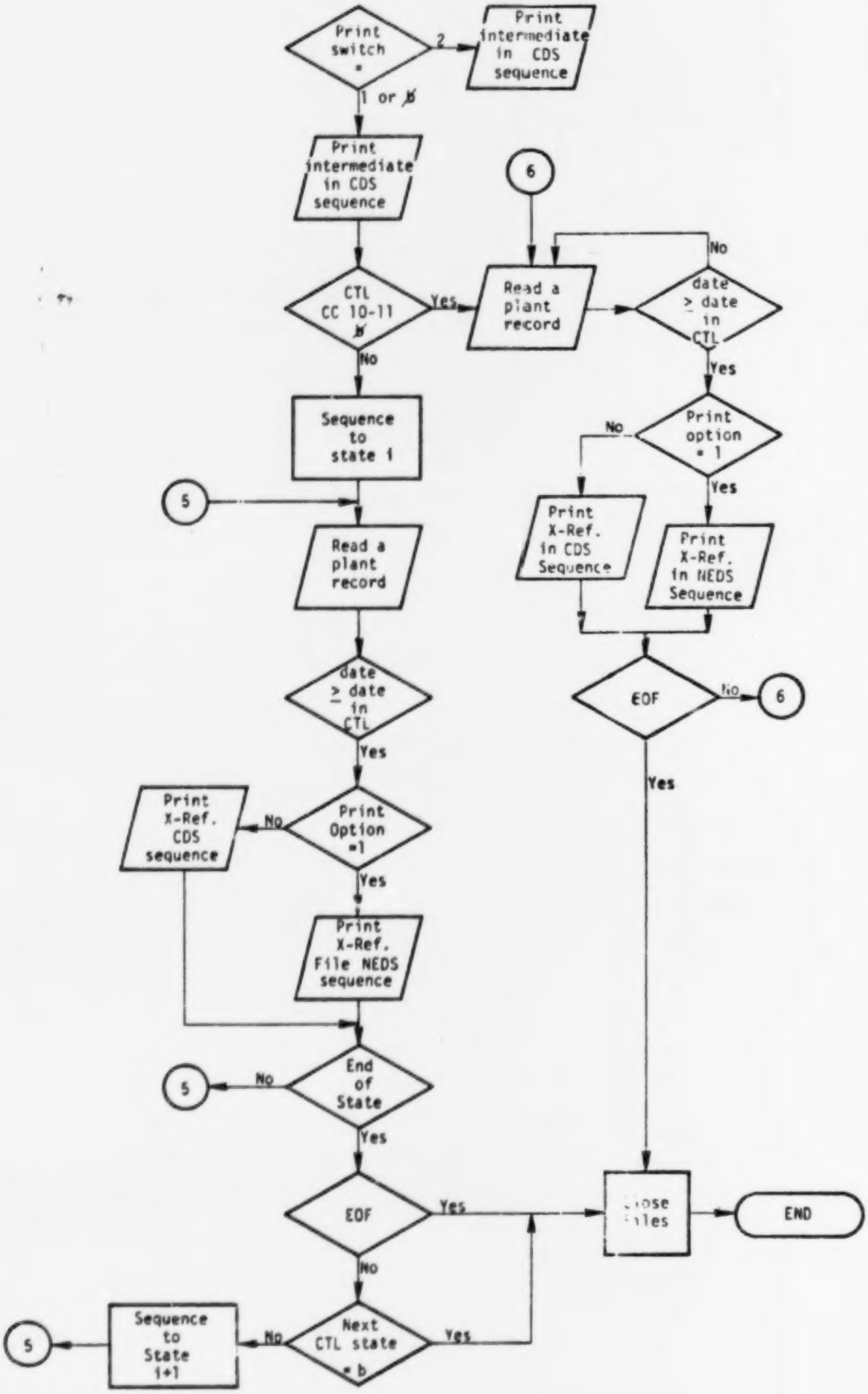
1.	Preliminary study = 160 hours @ $20/hr.	=	$3,200.00
	Clerical support = 40 hours @ $5/hr.	=	200.00
2.	Coordination with DSSE = 40 hours @ $20/hr.	=	800.00
3.	Program coding and debugging 120 hours@ $20/hr. plus 20 clerical support hours @ $5/hr.	=	2,500.00
4.	Program documentation = 32 hours @ $20/hr. plus 16 clerical support hours @ $5/hr.	=	720.00
5.	1 trip to DSSE = 2 days per diem @ $25/day plus air fare @ $60.00.	=	110.00
			$7,530.00
6.	Computer time is:		
	5 compile/tests @ 15 min/run = 75 min = 1.25 SUP @ $368/SUP	=	$ 460.00
			$7,990.00

PROGRAM LOGIC


START
Read CTL
Clear areas, set switches
2
Update switch =
2 or 3
Read CDS trans.
5
b
1
1 or 3
Read NEDS trans.
Read X-Ref. File
State, County, Plant I.D.'s =
No
Yes
4
Build Plant Arrays



End of Plant
No
Update switch =1
No
2
Yes
1
Yes
Update X-Ref. Plant
Write intermediate output
Clear work Arrays
EOF trans.
No
EOF X-Ref. File
No
Update switch =
1 or 3
1
2
2
Yes
Read Next Plant
4
Yes



Print switch =
2
Print intermediate in CDS sequence
1 or ƀ
Print intermediate in CDS sequence
6
CTL CC 10-11 ƀ
Yes
Read a plant record
No
date ≥ date in CTL
Yes
No
Print option = 1
Yes
Print X-Ref. in CDS Sequence
Print X-Ref. in NEDS Sequence
No
No
Sequence to state 1
5
Read a plant record
date ≥ date in CTL
Yes
EOF
No
6
Yes
Print X-Ref. CDS sequence
No
Print Option =1
Yes
Print X-Ref. File NEDS sequence
5
No
End of State
Yes
EOF
Yes
Close Files
END
No
Sequence to State i+1
5
No
Next CTL state = b
Yes

A.7

APPLICATION - Indicate sources that have been deleted from a report because of confidentiality.

ABSTRACT - This application will allow NEDS users to have a listing of Plant I.D.'s, Point Numbers, and Name and Address for all point sources deleted from a report because of confidentiality. The application requires a programming change to NE001A, the program that selects data from the NADB*NEDS-USER file and prepares them for sorting. The rationale for this change is that users may be able to better assess their reports if they know the degree to which the reports are complete.

INPUT DESCRIPTION - Input to the NE001A program will not change. The input is:

NADB*NEDS-USER file
NADB*NEDS-INDX-AQ file
NADB*NEDS-CNTY-ST file
NADB*NEDS-CNTY-AQ file
Control Card

OUTPUT DESCRIPTION - Output from this option will be a listing of the plant name and address, plant I.D. number, and point source numbers for plants having confidential SCC's. An example report format follows:

REPORT OF PLANTS WITH CONFIDENTIAL SCC NE 001A

STATE	COUNTY	AQCR	PLANT I.D.	NAME-ADDRESS (40 Char.)		POINT CODES
XX	XXXX	XXX	XXXX	X	X	XX
			XXXX	X	X	XX
			XXXX	X	X	XX
			XXXX	X	X	XX
			XXXX	X	X	XX
XX	XXXX	XXX	XXXX	X	X	XX
			XXXX	X	X	XX

PROGRAMMING REQUIREMENTS - The program logic in NE207A will need to be changed as follows:

1. The selection criteria should be changed to accumulate the following each time confidential SCC is encountered. The logic will be triggered by a code other than a "1" in CC 70 of the control card.
2. At the end of the job, the accumulated data will be printed in the format shown under "Output Description."

LEVEL OF EFFORT
and
ANTICIPATED COST -

1.	Preliminary study = 16 hours @ $20/hr.	=	$ 320.00
2.	Clerical support = 1 hour @ $5/hr.		5.00
3.	Program changes to NEOO1A = 16 hours @ $20/hr.	=	320.00
4.	Testing and debugging = 24 hours @ $20/hr.	=	480.00
5.	Documentation = 8 hours @$20/hr plus 2 hours @ $5/hr.	=	170.00
			$1295.00
6.	Computer time is :		
	2 compile/tests @ 15 min/run = 30 min = 0.5 SUP @ $368/SUP	=	$ 185.00
			$1480.00

A.8

APPLICATION - Implement user training seminars.

ABSTRACT - Two types of NEDS seminars should be offered:

1. Data preparation seminars
2. User-oriented seminars.

The data preparation seminars emphasize the techniques for coding and updating NEDS data. Emission Inventory Subsystem (EIS) compatibility with NEDS is included. A training manual has been developed, and several seminars have been held. An outline of the manual is included here, along with suggestions for revision. The format for the seminars basically follow the outline of the training manual.

The user-oriented seminars emphasize the applications for NEDS to agency operations. These seminars are directed more toward data systems and/or intermediate management levels than toward the technical personnel who might attend the data preparation seminars. A training manual and seminar agenda need to be developed.

DATA PREPARATION SEMINARS - The outline for the "National Air Data Branch Emission Data Systems Training Manual" is shown below.

ACKNOWLEDGMENT

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

2.0 NATIONAL EMISSIONS DATA SYSTEM (NEDS)

    2.1 Point Sources

2.1.1 Coding
2.1.2 Updating Point Source Data
2.1.3 Edit/Check Procedures

2.2 Area Sources

2.2.1 Data Development and Coding
2.2.2 Updating Area Source Data
2.2.3 Edit/Check Procedures

2.3 NEDS Output Formats

3.0 EMISSIONS INVENTORY SUBSYSTEM (EIS)

3.1 Synopsis of CDHS

3.2 Purpose and Capabilities of EIS

3.2.1 Purpose of EIS
3.2.2 Capabilities of EIS

3.3 NEDS-EIS Relationships

3.3.1 File Structures - NEDS/EIS
3.3.2 NEDS to EIS File Conversion
3.3.3 EIS to NEDS File Conversion

3.4 EIS Coding Requirements

3.4.1 EIS Coding Instructions
3.4.2 Add, Change, and Delete

REFERENCES

APPENDIX A SAROAD CODING CHANGES FOR NEDS PURPOSES

APPENDIX B STANDARD ABBREVIATIONS FOR USE IN COMMENTS

APPENDIX C TABLE FOR USE WHEN ASSIGNING PLANT IDENTIFICATION NUMBERS

APPENDIX D TEMPLATES OF NEDS POINT SOURCE CODING FORMS FOR USE WITH IBM FORTRAN CODING FORMS

The verification and validation procedures for NEDS should also be discussed in these seminars. The discussion would explain the formal procedures as outlined in the "AEROS User Manual." Suggestions should be made for helping state and local agencies as well as Regional Offices to follow the procedures within a reasonable time and with minimal impact on the agency. The required revision to Section 2.1 of the training manual is shown below.

2.1 Point Sources

2.1.1
2.1.2
2.1.3
2.1.4 Validation/Verification Procedures

The text in Section 2.1.4 will include:

2.1.4.1 Explanation of validation/verification procedures
2.1.4.2 Suggestions for following procedures with maximum efficiency.

USER-ORIENTED SEMINARS - An annotated outline for a training manual to be used in user-oriented seminars is shown below.

ACKNOWLEDGMENT

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION (Indicate that the purpose of the seminar is to acquaint managers with reports available from NEDS. Identify the topics to be discussed. The expected result is to increase the number of users.)

2.0 NATIONAL EMISSIONS DATA SYSTEM (NEDS)

2.1 Definition and Purpose of NEDS

2.2 Operating Procedures (data flow through the R.O.'s to NADB, etc.)

2.3 Status of NEDS (size of the data bank, completeness of data elements, representativeness in terms of total plant population, projected growth, etc.)

2.4 Available Reports

2.4.1 Report No.1 (Insert the report name here) (Include two subsections in the text for each report as follows:

2.4.1.1 Report format and retrieval options
2.4.1.2 Applicability for agency operations

2.4.2 Report No.2 (Repeat for each available report).

3.0 SPECIAL PROBLEMS

3.1 Trends Monitoring

3.2 Application to Air Quality Maintenance Area Plans

3.3 Application to strategy testing and modeling

4.0 OTHER SYSTEMS

(This section would briefly describe the other data bases related to NEDS, such as the Polk Vehicle File, FPC 67 File, and the Federal Facilities File.

The user-oriented seminars are to be directed to Regional Office personnel as well as to state and local agency personnel. The goal is to identify specific applications for NEDS within the Regional Offices and within the state and local agencies. Convincing potential users that NEDS can be used in their routine tasks would result in more enthusiastic participation in maintaining the data base. Before the user-oriented seminars could be implemented, NADB would have to perform a study to determine specific uses for emissions data within the three agency levels.

LEVEL OF EFFORT
and
ANTICIPATED COST - The costs that might be incurred with implementing each type of seminar are shown below. They are based on one seminar per year at each Regional Office.

DATA PREPARATION SEMINAR

1.	Revisions to training manual		
	40 professional man-hours/year @ $20/hr.	=	$ 800.00
	40 clerical support hours/year @ $5/hr.	=	200.00
			$1,000.00

2. Presentation of seminar

a. Manpower

10 trips 3 professional man-days/ trip, or 240 hours @ $20/hr.	=	$4,800.00
1 day followup per seminar, or 80 hours @ $20/hr.		1,600.00
1/2 day clerical support per trip, or 40 hours @ $5/hr.		200.00
		$6,600.00

b. Travel

30 days @ $25/day	=	750.00
Ground transportation @ $5/trip	=	50.00
Air Fare		
RDV - Chicago	=	132.00
- Boston	=	122.00
- New York	=	96.00
- Philadelphia	=	82.00
- Atlanta	=	84.00
- Kansas City	=	115.00
- Dallas	=	97.00
- San Francisco	=	378.00
- Denver	=	134.00
- Seattle	=	370.00
		$2,410.00
TOTAL		$10,010.00

USER-ORIENTED SEMINAR

1. Seminar development

Problem definition and preliminary investigation = 400 hours @ $20/hr.	=	$ 8,000.00
Seminar preparation = 80 hours @ $20/hr.	=	1,600.00
Manual preparation 80 hours writing and edit @ $20/hr. plus 120 hours clerical support @ $5/hr.	=	2,220.00
Printing costs for 500 copies @ 50 pages/copy @ 5¢/page		1,250.00
		$13,050.00

2. Seminar presentation These elements are assumed to be the same as those for the Data Preparation Seminar.	=	$10,010.00
Total (first year)		$23,060.00

A.9

APPLICATION - Add fugitive dust SCC's.

ABSTRACT - This application would result in new SCC's representing selected fugitive dust sources such as coke storage piles at steel plants. The recommendations and rationale for new SCC's should be a Regional Office responsiblity, since many of the industry types involved are regionally oriented. Because of the cost involved with adding new SCC's and developing new emission factors, preliminary studies should be conducted on the industries of interest. The preliminary studies should minimally address the following problems.

1. Current emission factor background information for the industry of interest must be investigated to ensure that the specific fugitive dust sources of interest have not already been included in the general emission factor.

   If it is in the emission factor, then that factor will need to be adjusted, if a separate factor for fugitive dust is desired.

2. A decision must be made as to whether the fugitive dust category could be considered a point source, or if it would better classify as an area source.

3. An estimate should be made of the probable impact of the fugitive dust source in relation to the total potential particulate emissions for several plants or processes representative of the industry.

LEVEL OF EFFORT - The factors determining the cost of adding new SCC's to the system are presented here. The Technical Data Section (TDS) of NADB must review the request for the new SCC. The tasks involved with adding a new SCC follow.

1. The current emission factors will have to be investigated to determine if the fugitive dust category has already been included in the factor. This will be an insignificant cost.

2. Estimating the relative contribution of the fugitive dust source to the potential emissions of the associated process or processes will incur a relatively insignificant cost, since source testing, etc. will already have been completed for most of the sources of interest.

3. Emission factors will need to be developed if SCC's are added. The development of emission factors will be the most significant cost associated with adding SCC's. Factor development will require:

    a. Literature search of similar processes or sources.

    b. Correlation of properties of the source in question with those of similar sources.

    c. Correlation with average operating parameters, i.e. seasonal considerations, daily process fluctuations, etc. to establish the SCC units.

    d. Ambient monitoring to determine relative contributions of the fugitive dust source. This requirement is optional, depending on the industry and source type being investigated, and it is not included in the cost estimates shown below.

4. If a general SCC has already been assigned for the industry, the emission factor associated with that SCC may have to be changed to reflect the contribution from the new SCC.

5. Expanded instructions and changes to existing process descriptions in AP-42, Compilation of Air Pollutant Emission Factors will be required.

6. The NEDS records that show general SCC's, i.e. 99's, will have to be reviewed and updated for the new SCC and revised emissions estimates.

The table below shows the professional and clerical support man-hours expected for each task.

Table A.9.1

SUMMARY OF LEVEL OF EFFORT FOR ADDING AN SCC

Task	Professional hours	Clerical support hours
1	16/emission factor	
2	40/fugitive dust source	
3a	120	40
3b	40	
3c	40	
4	10	
5	80	40
6	200	40
TOTAL	546	120

546 Professional hours @ $20/hr. = $10,920.00

120 Clerical support hours @ $5/hr. = 600.00

TOTAL = $11,520.00/SCC

A.10

APPLICATION - Improve the format for identifying report retrieval options.

ABSTRACT - The purpose of this application is to provide NEDS users with an easy reference to determine the type of reports available and the retrieval options. A general publication should include:

1. example report formats
2. abstract of the report
3. tabular cross-reference for identifying retrieval options

Examples already exist within some Regional Offices, and no significant impact on NADB's operating procedures or budget is expected. The following table is an example of the matrix for defining retrieval options.

LEVEL OF EFFORT
and
ANTICIPATED COST - Total effort for this application is minimal, since the necessary materials already exist. Total professional time and clerical support are not anticipated to exceed:

1. Professional hours = 40 @ $20/hr.	=	$ 800.00
2. Clerical support = 40 @ $5/hr.	=	200.00
		$1,000.00
Report reproduction = 500 copies 20 pages/ report @ 5¢/page.		$ 500.00
	TOTAL	$1,500.00

NEDS DATA SYSTEM

KEY WORD	POINT SOURCE	CONDENSED POINT SOURCE	AREA SOURCE	AREA SOURCE FUEL USAGE	SOURCE COUNTING	ALLOWED VERSUS COMPUTED EMISSION	EMISSION SUMMARY	STATIONARY SOURCE FUEL SUMMARY
EPA Region								
State								
State/County								
State/County/Plant								
State/Co./Plant/Point								
State/SCC								
State/County/SCC								
Ownership Code								
SIC								
Method of Estimation								
SCC								
AQCR								
Sort Options								

(Indicate by "yes" or "no" retrievals that are available)

APPENDIX B

TASK DESCRIPTIONS
SAROAD

APPENDIX B TASK DESCRIPTIONS - SAROAD

B.1.

APPLICATION - Trends plotting.

ABSTRACT - This application will allow graphic presentation of arithmetic and geometric means for a pollutant at a specified site over an extended period of time. A total plotting package identified in the survey would include the capacity to produce five types of plots:

1. High 8-hour CO ave/wk/yr.
2. High 24-hour SO_2 ave/wk/yr.
3. Monthly averages/any pollutant
4. Yearly averages/any pollutant
5. Percent of days/time period the station is in violation of the standard.

The specifications written here are for a generalized approach to modifying existing programs or for writing new programs to produce the data in required format for plotting averages. The plotting will be done on-line using the SAS package.

INPUT DESCRIPTION - For most applications, the input will come from one of the following files:

1. NADB* NADB-YRSUM-D
2. NADB* NADB-QRSUM-D
3. NADB* NADB-MOSUM-D

The application file will be read and the necessary data written out as punched cards or as card image format magnetic tape. The file will be searched by the following key sequences:

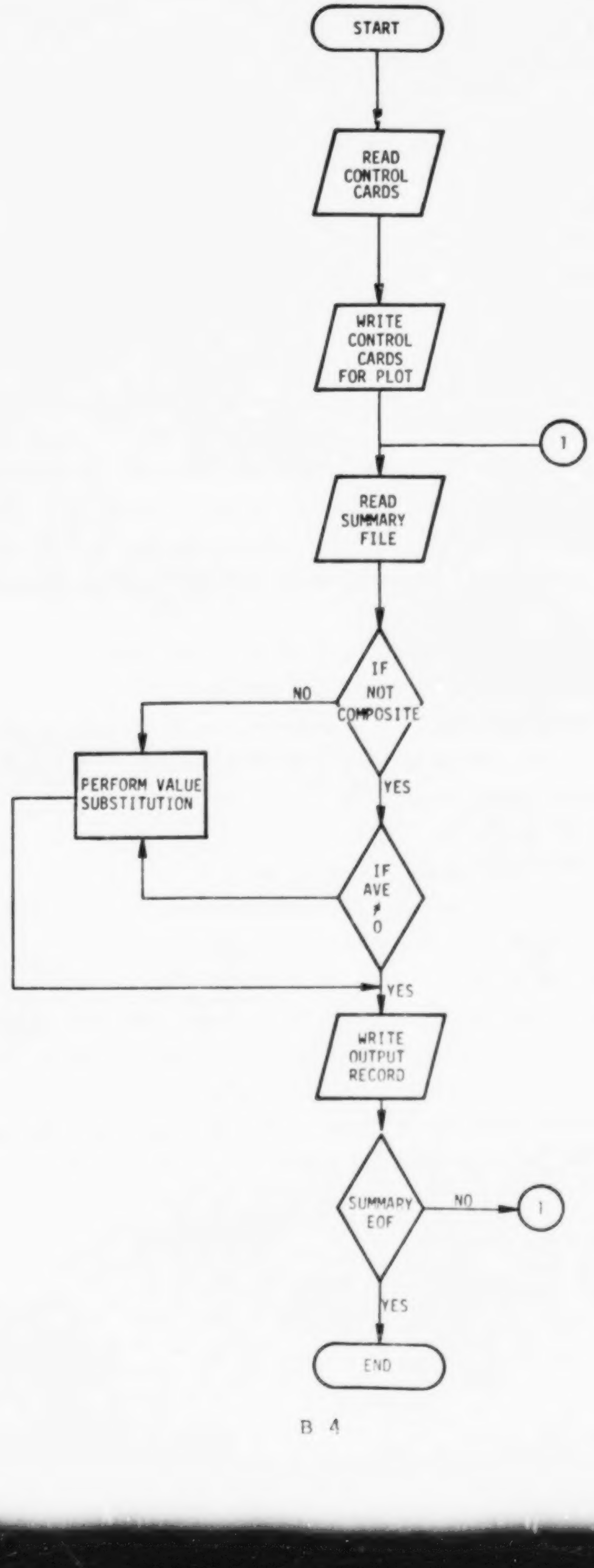
State
Area
Site
Agency
Project
Pollutant
Method
Interval
Year

OUTPUT DESCRIPTION - Output will be punched cards or card - image tape suitable for input to a plot routine. The output records will include the data values to be plotted. The data values will be preceded by sufficient identifier cards to:

1. Set up the scale for the plot routine
2. Provide the necessary graphic labelling information, i.e. assign scale values and label the graph as "Co 8-hour values for site xxx", etc

PROGRAM LOGIC - A conceptual logic flow for inclusion in new programs or in current analysis programs that might be modified is shown in the flow diagram.

A subroutine will be needed to handle the special case in which not enough data have been reported for a period to calculate the desired statistics. This subroutine will also need to handle cases in which the values in the summary record represent composite values,i.e. values for periods other than monthly, quarterly, or yearly.


START
READ CONTROL CARDS
WRITE CONTROL CARDS FOR PLOT
1
READ SUMMARY FILE
IF NOT COMPOSITE
NO
YES
PERFORM VALUE SUBSTITUTION
IF AVE ≠ 0
YES
WRITE OUTPUT RECORD
SUMMARY EOF
NO
1
YES
END

LEVEL OF EFFORT
and
ANTICIPATED COST - (Based on six applications)

1.	Preliminary study and definition of initial package specifications = 120 hours @ $20/hr.	=	$2400.00
2.	Changing current retrieval programs or developing new program/programs = 120 hours @ $20/hr. plus 32 hours clerical support @ $5/hr.	=	2560.00
3.	Testing and debugging retrieval programs = 80 hours @ $20/hr.	=	1600.00
4.	Writing and testing plot routine specifications 40 hours @ $20/hr.	=	800.00
5.	Documentation = 40 hours @ $20/hr plus 40 hours clerical support @ $5/hr.	=	1000.00
			$8360.00
6.	Computer time is:		
	(1) 12 compile/tests @ 15 min/run for retrieval programs = 180 min = 3 SUP @ $368/SUP.	=	$1105.00
	(2) 12 tests for plots @ 10 min/test = 120 min = 2 hrs @ $50/hour (assumes that cost of plotting would not exceed $50/hour)	=	$ 100.00
			$9565.00

B.2.

Application - Capability to audit SIP stations reporting.

ABSTRACT - The purpose of this program is to provide the Regional Offices with the capability to track data reported by the states for SIP required monitoring stations. The program is in COBOL. A tape file of SIP required instrumentation is generated, and it is compared to the sorted input tape to the SAROAD editor program. A listing of instruments with no data submitted or of instruments not meeting the 75% criterion is produced. The 75% criteria logic is incorporated in the program. This program will save most Regional Offices at least 2 1/2 man-days per state per quarter.

INPUT DESCRIPTION - The input to this program is the sorted card image transaction tape of SAROAD update transactions before edit, a tape file of SIP - required station numbers, and an audit tape.

The format of the SIP Station Numbers Tape record follows:

Position	Description	Format
1-2	State Code	99
3-6	AQCR	9(4)
7-9	Site Code	9(3)
10	Agency Code	X
11-12	Project Code	99
13-17	Pollutant Code	9(4)
18-19	Method Code	99
20	Interval	X

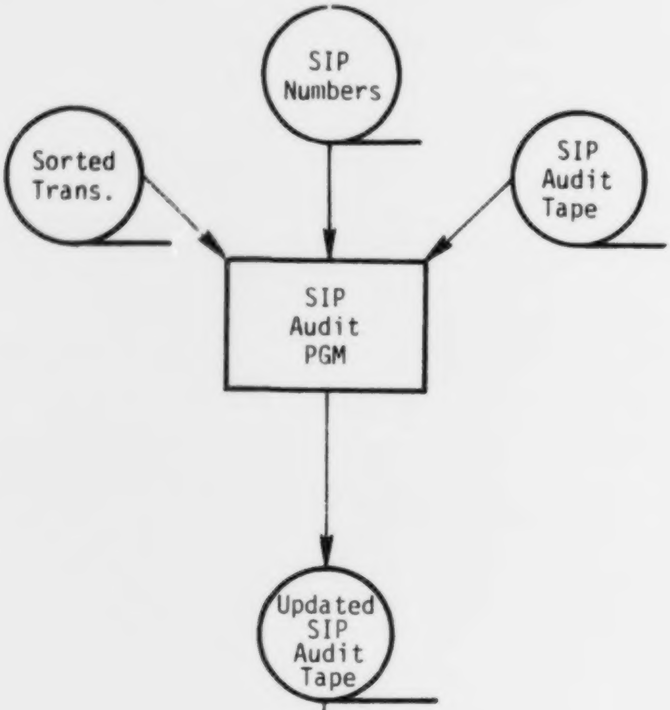
OUTPUT DESCRIPTION - The output to this program is a report listing the Master Key for each SIP instrument within a Regional Office for which no data was reported during the quarter or for which the 75% criterion was not met. The format example is shown in Figure B.2.1. The control break is on Region number.

SYSTEM FLOW - This job requires that the SIP Audit Tape be updated quarterly to add SIP sites that have been added to SAROAD or to change the keys for any sites at which the Agency/Project codes may have changed during the quarter. The mechanism for identifying a change to the Agency/Project code and the tracing of the change through the system has not been finalized. This mechanism will help to define the method to be used to update the Audit Tape. An alternative update method would require the Regional Office to submit a change card to the Audit Program when the Agency/Project code changed or when a site was added to or deleted from SAROAD as a SIP site. The specifications here are based on this method of updating. It will also be necessary to program the necessary logic to update the SIP Station Numbers Tape.

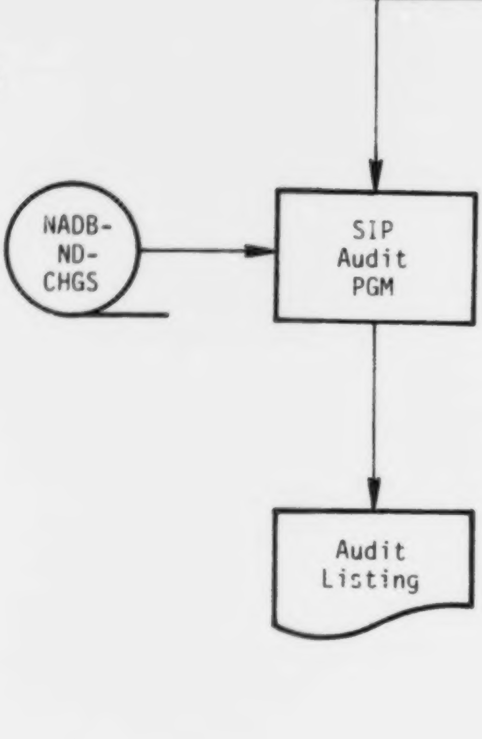
```
                 AUDIT OF SIP INSTRUMENTS - - - - - - REGION

YEAR     QUARTER   DATE OF RUN
 XX         X       XX  XX  XX

STATE   AQCR   SITE   AGENCY   PROJ.   POL.   MTHD.   INT.   NO  DATA   L.T. 75
 XX     XXXX   XXX      X       XX     XXXX    XX      X         X
                                       XXXX    XX      X                  X
                                       XXXX    XX      X         X

               XXX      X       XX     XXXX    XX      X                  X
 XX     XXXX   XXX      X       XX     XXXX    XX      X         X
        XXXX   XXX      X       XX     XXXX    XX      X         X
```

Figure B.2.1 Example of SIP Audit Report.

STEP 1 - (Update the SIP Audit Tape)


SIP Numbers
Sorted Trans.
SIP Audit Tape
SIP Audit PGM
Updated SIP Audit Tape

STEP 2 - PRINT


NADB-ND-CHGS
SIP Audit PGM
Audit Listing

LEVEL OF EFFORT
and
ANTICIPATED COST -

1.	Preliminary investigation and problem definition = 40 hours @ $20/hr.	=	$ 800.00
2.	Program coding = 120 hours @ $20/hr. plus 20 hours clerical support @ $5/hr.	=	2,500.00
3.	Testing and debugging = 60 hours @ $20/hr.	=	1,200.00
4.	Documentation = 32 hours @ $20/hr. plus 16 hours clerical support @ $5/hr.	=	720.00
			5,220.00
5.	Computer time is:		
	6 compile/tests @ 20 min/run = 120 min = 2 SUP @ $368/SUP	=	$ 735.00
			$5,955.00

B.3.

APPLICATION - Graphics to plot site locations.

ABSTRACT - This program will allow graphic display of the relative locations of sampling sites within a polygon area specified by latitude/longitude coordinates. User input is a map scaling factor, latitude/longitude coordinates defining the boundaries of the polygon, and SAROAD geographical retrieval keys. The program reads the SITE file and determines which stations have coordinates that lie within the defined polygon. Output is punched cards or card-image magnetic tape for plotting with a standard plotter. The plotted output shows the polygon, and the relative location of each station. The site number is printed beside an X marking each site location. The program is written in FORTRAN. The program operates in remote batch mode.

INPUT DESCRIPTION - Input will be punched card input or keyboard input in remote batch mode. The State, and County numbers will be input with the coordinates that define the area. A control card will be used to define a parameter "n", the number of vertices (points) on the polygon. The control card will also define the scaling factor to be used in the plot. The SITE file is included as input. The format of the input cards is:

Card 1

CC	Description	Format
1-2	State I.D.	I2
3-5	AQCR 1	I3
6-9	County I.D. 1	I4
. . .	(Repeat County I.D. for up to 10 counties/card)	. . .
80	"1"	I1

Card 2

CC	Description	Format
1-2	Number of points on the polygon.	I2
3-9	Longitude coordinate for first point	F7.0
10-15	Latitude coordinate for first point	F6.0
. . .	Repeat above format for each polygon point up to 6 points/card	. . .
75-80	Scaling factor	F6.0

OUTPUT DESCRIPTION - Output will be card image format records for use on a plotter. The first card defines the scaling factor and the polygon. One card per sampling station within the polygon will be output. The polygon will be traced, and the sites marked at their respective locations. The output card format is:

Card 1 (Defines the Polygon)

CC	Description	Format
1-6	Scaling Factor	F6.0
7-13	Longitude coordinate for first polygon point.	F7.0

CC	Description	Format
14-19	Latitude coordinate for first polygon point	F6.0
.		
.	Repeated for each point up to 6 points/card	
.		

Card 2 (One card for each site)

CC	Description	Format
1-3	Site No.	I3
4	blank	X
11-17	Longitude coordinate	F7.0
18	blank	
19-24	Latitude coordinate	F6.0

OPERATION AND PROGRAM LOGIC - The operating restrictions are the same as those described in Section A.2, Polygon-Defined Area Retrieval. The polygon can be any triangle, any rectangle, or any convex polygon, i.e. one for which the angle between any two consecutive sides is greater than 90 degrees but less than 180 degrees. Any polygon with five or more sides having an angle less than 90 degrees between any two consecutive sides is unacceptable. The programming requirements and logic are the same as those in Section A.2 with the following exceptions:

1. Selection keys are: State, AQCR, County, or combinations of these.
2. A scaling factor, corresponding to the input scaling factor, is punched as part of the output.
3. The coordinates of the stations of interest are retrieved from the SITE FILE.

LEVEL OF EFFORT
and
ANTICIPATED COST -

1. Preliminary investigation and problem definition = 40 hours @ $20/hr. = $ 800.00

2.	Coding retrieval program = 80 hrs @ $20/hr. plus 16 hours clerical support @ $5/hr.	=	1,680.00
3.	Testing and debugging = 80 hours @ $20/hr.	=	1,600.00
4.	Testing and debugging = plot routine = 16 hours @ $20/hr.	=	320.00
5.	Documentation = 24 hours @ $20/hr. plus 16 hours clerical support @ $5/hr.	=	560.00
			$4,960.00
6.	Computer time is:		
	1) Assume same time as polygon retrieval	=	$1,475.00
	2) 3 test runs of plot routine @ 10 min/run = 1/2 hour @ $50/hr.	=	25.00
			$6,460.00

B.4

APPLICATION - Report on standards violations.

ABSTRACT - This application allows users to see only those statistics necessary to compare national air quality standards with pollutant levels measured at a specified site for each of the criteria pollutants. This report shows primarily the same standards data as the Standards Report, but it excludes the individual values for each interval. The program is in COBOL.

INPUT DESCRIPTION - Input to this program is the NADB* NADB-ND-SITE file and the Standards tapes. Selection is by State, AQCR, pollutant, year, or any combination of these.

OUTPUT DESCRIPTION - The output is a printed report showing the following data:

1. Site Code
2. Agency/Project Code
3. AQCR Code
4. Pollutant/Method Code
5. Interval
6. Year
7. Number of violations for year
8. Average concentration at this site for the pollutant/method/interval code
9. Standard for the pollutant interval
10. Code to show if standard is primary or secondary
11. Percent of time standard was exceeded

PROGRAMMING REQUIREMENTS - Only those site/pollutant combinations that have exceeded the standards will be printed. For each record processed, an algorithm must be applied to compute the mean value for the site/pollutant/method if a violation has occurred within the year.

LEVEL OF EFFORT
and
ANTICIPATED COST -

1.	Preliminary investigation and problem definition = 40 hours @ $20/hr.	=	$ 800.00
2.	Coding = 40 hours @ $20/hr plus 16 hours clerical support @ $5/hr.	=	880.00
3.	Testing and debugging = 40 hours @ $20/hr.	=	800.00
4.	Documentation = 16 hours @ $20/hr plus 8 hours clerical support @ $5/hr.	=	360.00
			$2,840.00
5.	Computer time is:		
	3 compile/tests @ 15 min/run = 45 min = .75 SUP @ $368/SUP.	=	$ 275.00
			$3,115.00

B.5

APPLICATION - Include data not meeting the 75 percent criteria in the summary reports.

ABSTRACT - This application gives SAROAD users the option to include data not meeting the 75 percent criteria in summary reports. The logic of the reporting programs will be altered to the extent that when the criteria flag is not set in the summary file statistics for the associated record will be calculated and included in the report to indicate that the statistics shown were calculated from data that did not meet the 75 percent criteria. The programs affected are NA202, which produces the Inventory Report, and NA212, which produces the Yearly Report by Quarters.

INPUT DESCRIPTION - The NADB* NADB-QTRSUM- file and the NADB* NADB-YRSUM-D file are input to NA212. The NADB*NADB-YRSUM-D file is input to NA202. The record descriptions for the files are identical.

OUTPUT DESCRIPTION - The output from both NA202 and NA212 will include a flag to indicate cases in which reported statistics include data not meeting the 75 percent criteria.

PROGRAM SPECIFICATIONS -

1. The NA202 program needs to be changed to calculate the arithmetic mean for records in the NADB*NADB-YRSUM-D file that have blanks in the criteria field. The arithmetic mean is calculated by:

$$\text{arithmetic mean} = \frac{\text{sum } x}{\text{No. of observations}}$$

The computed arithmetic mean is not added into the file.

2a. The NA212 program needs to be changed to calculate an arithmetic mean for each quarter in which the criteria flag is blank. The quarter number (position 23-24) must be interrogated to associate the calculated result with the proper print field. The arithmetic mean is calculated by:

$$\text{arithmetic mean} = \frac{\text{sum x}}{\text{No. of observations}}$$

The computed arithmetic mean is not added into the file.

b. For records that are composites rather than quarterly data, the report should be flagged to indicate that the calculated statistics represent a composite.

c. When the criteria flag is blank in a record in the NADB*NADB-YRSUM-D file, the arithmetic mean and arithmetic standard deviation should be calculated. The arithmetic mean is caluclated as:

$$\text{arithmetic mean} = \frac{\text{sum x}}{\text{No. of observations}}$$

The arithmetic standard deviation is calculated as:

$X = \text{sum } x^2 - \text{SUMX} * \text{SUMX/No. of observations}$

standard deviation = square root of X/(No. of obser. -1)

A subroutine to calculate the square root is required.

3. For both programs, a flag should be included on the report to indicate statistics that have been calculated using data that do not meet the 75 percent criteria.

INPUT FILE FORMAT

Position	Description	Picture
1-24	Key	X(24)
25-28	% of observations	9(4)
29-32	arithmetic mean	9(4)
33-36	log-mean	9(4)
37-40	geometric mean	9(4)
41-44	arithmetic standard deviation	9(4)
45-48	log-standard deviation	9(4)
49-52	geometric standard deviation	9(4)
53-56	2nd moment	9(4)
57-60	max value	9(4)
61-64	med value	9(4)
65-68	min value	9(4)
69-72	zero sub	9(4)
73-74	No. of observations	99
75-76	substitutions made	99
77-80	sum x	9(4)
81-84	sum x^2	9(4)
85-88	sum LN(x)	9(4)
89-92	sum $LN(x)^2$	9(4)
93-94	units	99
95	blank	x
96	criteria flag	x
97-140	blank	x(43)

LEVEL OF EFFORT
and
ANTICIPATED COST - The following estimates cover changes to both programs.

1.	Preliminary investigation and problem definition = 40 hours @ $20/hr.	=	$ 800.00
2.	Coding the program logic = 30 hours @ $20/hr plus 8 hours clerical support @ $5/hr.	=	640.00
3.	Testing and debugging = 50 hours @ $20/hr.	=	1000.00
4.	Documentation = 8 hours @ $20/hr plus 8 hours clerical support @ $5/hr.	=	200.00
			$2640.00
5.	Computer time is:		
	2 compile/tests per program @ 10 min/run = 40 min = .75 SUP @ $368/SUP.	=	275.00
			$2915.00

B.6

APPLICATION - Include 2nd maximum measured value on the Standards Report.

ABSTRACT - This application allows users to see the second maxiumum measured value for a pollutant/interval at a specific site on the Standards Report for that pollutant/ interval. Implementation requires the accumulation of the second maximum value for each month and the addition of a print line at the bottom of the report to show the second maximum values for all months in the year of interest. The following programs require modification:

1. NA224-NA234
2. NA251
3. AE003

INPUT/OUTPUT DESCRIPTION - Input to NA224-NA234 is:

1. AERO-AOCR-NM
2. AERO-CTYCNTY
3. NADB-STE-INX
4. AERO-SMSA-NM
5. NADB-ND-SITE
6. ND subfile created from NA235

Output is the current Standards Report with an extra line at the bottom to show second maximum values.

PROGRAMMING REQUIREMENTS - No file changes are necessary. Programs NA224-NA234 must be modified as follows:

1. Format a print line for the 2nd max. values
2. Insert logic to determine the 2nd max. value for each month and move the value to the proper field in the new print line.

The same basic logic as applies to determining the maximum values can be applied for determining 2nd max. values with slight modification.

Programs NA251 and AE003 will be modified to handle the extra print line.

LEVEL OF EFFORT and ANTICIPATED COST -

1.	Preliminary investigation and problem definition = 30 hours @ $20/hr.	= $ 600.00
2.	Coding = 24 hours @ $20/hr plus 4 hours clerical support @ $5/hr.	= 500.00
3.	Testing and debugging = 16 hours @ $20/hr.	= 320.00
4.	Documentation = 8 hours @ $20/hr plus 4 hours clerical support @ $5/hr.	= 180.00
		$1,600.00
5.	Computer time is:	
	3 compile/tests @ 10 min/run = 30 min = 0.5 SUP @ $368/SUP	= $ 185.00
		$1,785.00

B.7

APPLICATION - Identify inactive sites.

ABSTRACT - This application will add an active/inactive code to each record in the Site File. Major system changes are required in order to make this a useful feature. In order to adequately trace changes in a site code throughout the system, four site status codes are required:

1. Active site
2. Inactive site
3. Modified active site
4. Modified inactive site

Applicable dates for site number changes, etc., will also be maintained within the Site Description File. The rationale for this application is that it will allow users to perform trends analysis using the computer. The specifications included here are primarily those suggested by the Monitoring and Reports Branch (MRB) of the Monitoring and Data Analysis Division.

INPUT DESCRIPTION - Input to this application is the SAROAD Site Identification form. The form should be revised as necessary to allow inclusion of the site number previously used. A collection code has also been recommended. This code will be used to flag cases in which sample collection and sample analysis are performed by different agencies.

OUTPUT DESCRIPTION - The output from this application will be a reformatted NADB* NADB-STE-D file (Site Description File).

LOGIC - The following is a summary of the recommendations made by MRB for changing the Site Identification Record:

1. Make the Parameter Code part of the site identification
2. Add a status field to show that the site is no longer active. Purge inactive sites from the Site File after five years.
3. Add an "active date" field for each site to indicate the date that the site became active.
4. Include a "terminal date" field for sites that become inactive or are modified significantly enough to be considered new sites.
5. A "related site" field would be added to each new site record (if the site replaces a previous site) to indicate the previous number.
6. A "collection code" field to indicate the type of agency collecting the data (as opposed to the type of agency performing the analysis).

MRB suggests that all air quality data for a site be changed to reflect the new active site code whenever a site is modified (in accordance with SAROAD Users Manual specifications, i.e. when the monitor location has been moved to such a degree that equivalent ambient air is not being monitored). This would be done to ensure that all ambient air data collected for the same location under a like environment are stored together and are retrieved easily. Figure 1 illustrates the relationship between data and site information per MRB's recommendations.

1. Site 152040010A01 Becomes Operational 5/1/68
2. Site Changed to 152040010H01 6/1/72
3. Site Changed to 152040010F01 7/1/73

	DATE		SITE DESCRIPTION FILE					A.O. DATA	
		SITE	POL.CD.	ST.CD.	COLL.CD.	ACTIVE DATE	TERM DATE	REL SITE	SITE
1.	01/01/69	152040010A01	11101	A	A	05/01/68			152040010A01
	05/30/72	152040010A01	11101	A	H	05/01/68			152040010A01
2.	08/01/72	152040010A01	11101	B	H	05/01/68	06/01/72		
		152040010H01	11101	A	H	06/01/72		A01	152040010H01
3.	09/06/74	152040010A01	11101	B	H	05/01/68	06/01/72		
		152040010H01	11101	B	H	06/01/72	07/01/73	A01	
		152040010F01	11101	A	F	07/01/73		H01	152040010F01

(This figure is reproduced from a description obtained from MRB during the survey interview.)

PROJECTED IMPACT - Implementation of this approach to storing site description information will require reformatting the NADB*NADB-STE-D file (Site Description File). The following files are created from the Site Description File, and they will require the same reformatting:

1. NADB*NADB-STE-T (tape backup to disk file)
2. NADB*NADB-STE-INX (control file for edit process)
3. NADB*NADB-STE-LST (tape file of print line images of sampling site descriptions).

The following files will need to be reformatted to reflect the addition of the "collection code".

1. NADB*NADB-YRSUM-D (for Inventory Listing from NA202).
2. NADB*NADB-ND-SITE (for row data listings).

If inclusion of the "collection code" in the headings of the Inventory Listing or any of the raw data listings is not necessary, these two files will not require reformatting, nor will their associated programs require any changes.

If pollutant code (probably pollutant/method code) is made part of the site identifcation, then the NADB*NADB-STE-D file will need to be expanded. Currently, the file has approximately 10,500 site description records. The file would be expanded to a minimum size of approximately 32,000 site description records, if it is assumed that an average of three pollutant method codes exist for each site in the file (from brief examination of the report titled "Status of the National Aerometric Data Bank NADB as of November, 1973"). Compiling the necessary information to reformat the file would require limited special programming to list site/pollutant/method codes so that they can be reviewed before the Site File is changed.

Including the pollutant/method code as part of the site description will necessitate minor changes to most report-generating programs in the SAROAD system.
The total impact of changing the site codes in SAROAD necessitates a comprehensive review of the total system.

B.8

APPLICATION - County retrieval of data

ABSTRACT-- This application will allow users to retrieve reports for sites within county. No new codes are necessary, since a county code is already part of the site identification record. No new files are necessary since the NADB*AERO - AQCR file (contains state/county numbers within AQCR) and the NADB*AERO-STEAQCR file (contains site codes within AQCR) together contain the necessary sort key information. Each report program using the county retrieval option will have to be modified to incorporate these sort keys.

INPUT-OUTPUT DESCRIPTION - The input and output will depend on the individual programs selected to include county retrieval.

PROGRAMMING REQUIREMENTS - Each program will require the incorporation of the necessary logic to use the NADB* AERO-AQCR file and the NADB* AERO-STEAQCR file to allow report retrieval by coordinated record keys.

LEVEL OF EFFORT
and
ANTICIPATED COST -

1.	Preliminary investigation and problem definition = 60 hours @ $20/hr.	=	$1,200.00
2.	Coding = 24 hours @ $20/hr plus 4 hours clerical support @ $5/hr.	=	500.00
3.	Testing and debugging = 16 hours @ $20/hr.	=	320.00
4.	Documentation = 4 hours @ $20/hr plus 4 hours clerical support @ $5/hr.	=	100.00
			$2,120.00

5. Computer time is:

   2 compile/tests @ 15 min/run = 30 min = $ 185.00
   = 0.5 SUP @ $368/SUP

   $2,305.00

B.9

APPLICATION - Include minimum detectable levels on reports.

ABSTRACT - This application allows users to see the minimum detectable level for a specific pollutant/method on the following reports:

1. Yearly Report by Quarters
2. Yearly Frequency Report
3. Quarterly Frequency Report
4. Raw Data (24 hour) Listings

The estimate of minimum detectable levels for each pollutant/method are in the NADB*NADB-PARMFL. Programs that require modification to include this information in their output are:

1. NA211
2. NA212
3. NA213
4. NA219

(Modification to standards reports would require changes to a long series of programs, and should be carefully considered by NADB).

INPUT/OUTPUT DESCRIPTION - All four of the above programs use the NADB*NADB-PARMFL. The programs NA211-213 will be modified to include a fifth line of descriptive information for each pollutant/method. NA219 will be modified to include the minimum detectable level at the bottom of the page.

PROGRAMMING REQUIREMENTS - The output formats for each program will be modified to include an extra print line for each pollutant/method. Changes for NA211-NA213 will be the same for each program.

LEVEL OF EFFORT
and
ANTICIPATED COST - (4 programs)

1.	Preliminary investigation and problem definition = 16 hours @ $20/hr.	=	$ 320.00
2.	Coding = 16 hours @ $20/hr plus 2 hours clerical support @ $5/hr.	=	330.00
3.	Testing and debugging = 16 hours @ $20/hr.	=	320.00
4.	Documentation = 4 hours @ $20/hr plus 2 hours clerical support @ $5/hr.	=	90.00
			$1,060.00
5.	Computer time is:		
	4 compile/tests @ 15 min/run = 60 min = 1 SUP @ $368/SUP.	=	$ 370.00
			$1,430.00

APPENDIX C

INTERVIEW SUMMARIES

APPENDIX C INTERVIEW SUMMARIES

SURVEY PARTICIPANTS

1. EPA REGION I - Boston

Tom Devine
John Courcier
Bill Servoy

Don White
Val Deschamps
Jerry Levy

- Needham

Bill Walsh
Richard Rogers
Norm Beloin
Alberto Costales

Arnie Leriche
John Feldman
Dave Stonefield
Joe Mercadante

2. EPA REGION II

Gerard Sofian
Alex Salpeter
Ed Gawlinski

3. EPA REGION III

Dan Fitzgerald
Ed Skernolis

William Belanger
Connie Carr

4. EPA REGION IV

Barry Gilbert
Tommie Gibbs
Mike De Busschere

Carolyn Heller
Vince Helwig
Ray Cunningham

5. EPA REGION V

Dr. R. Trautner
Dr. P. Cho
B. Bolka
S. Goranson
M. Dipert
P. Gillen
T. Voltaggio

L. Larsen
L. Lehrman
Dr. B. Fairless
B. Kramer
D. Hoglund
R. Van Mersbergen
J. Logsdon

6. EPA REGION VI

E. Ray Lozano
Chris Jacobs
Marvin Waters
Carl Townsend
Peggy Reiff

Kay Dove
Jeannean Hayes
Stanley Spruiell
Mary Marusak
Doug Grano

7. EPA REGION VII

Art Spratlin
Charles Whitmore

Seymour Shuster
Michael Anderson

8. EPA REGION VIII

Jim Harris
John Dion
Doug Skie
Bob Fackler
Dale Wells

9. EPA REGION IX

Steve Body
Mark Brucker
Greg Fischer
Mike Stenburg
Jim Grove
Charlotte Hopper
Rob Ireson
Lloyd Kostow
Carolyn Lewis

10. EPA REGION X

A. E. Parlier
Kenneth Feigner
Shirley Schmidt
Mike Anderson
George Hofer
Cecil Drotts

11. GRANTS INFORMATION BRANCH OF GRANTS ADMINISTRATION DIVISION

Paul Wagner

12. OFFICE OF PLANNING AND EVALUATION

James Janis
Frank Blair

13. COMPLIANCE AND ANALYSIS SECTION, DIVISION OF STATIONARY SOURCE ENFORCEMENT

Michael Merrick
Carl Edlund

14. CONTROL SYSTEMS LABORATORY - NERC/RTP

Charles Chatlynne
James Wingo

15. HUMAN STUDIES LABORATORY - NERC/RTP

Dr. Bill Nelson
Vic Hasselblad

16. STRATEGIES AND AIR STANDARDS DIVISION - DURHAM

Justice Manning
Dennis Ludwig
Ray Morrison
Dick Atherton

17. NATIONAL AIR DATA BRANCH - DURHAM

Jim Southerland
Arch McQueen
Chuck Mann

18. OFFICE OF AIR AND WASTE MANAGEMENT

Lazlo Bockh

19. MONITORING AND DATA ANALYSIS DIVISION

Dr. Tom Curran
Bob Faoro
Jon Clark
George Manire
Neil Berg
Tom McMullen
Virginia Henderson
Marty Martinez

20. NATIONAL COUNCIL ON ENVIRONMENTAL QUALITY

Dr. James Reisa

PRE-SURVEY INFORMATION

The following letter and the 11 pages following it was sent to the NEDS/SAROAD coordinators in each Regional Office prior to the interviews.

"The National Air Data Branch, in an effort to be responsive to the practical needs of EPA users and potential users of the EPA air data bases, has contracted with PEDCo-Environmental Specialists, Inc. to conduct a survey among EPA Regional Offices and other selected users of the data bases. The purpose of the survey is to define current uses of the data bases and to project future user requirements. Comments on current system capabilities regarding operating procedures, report formats, storage and retrieval options, and summary capabilities will be required to make this survey meaningful. Suggestions for improvement are an important part of the survey.

The contractor will summarize the results of the survey, and provide NADB with suggestions for implementing user requirements. All users of the data bases should participate in the survey. This is an opportunity for users to provide significant input to the future development plans for the data bases. The users should feel free to be candid with the contractor's representatives. A handout of materials for distribution to personnel is enclosed. This hopefully will provide personnel with some orientation to the kinds of subjects to be discussed. Discussions need not be restricted to these items.

PEDCo will contact the Regional Office NEDS/SAROAD coordinators to arrange a schedule for visitation. In most cases, PEDCo personnel will be available for interviews for two days at each Regional Office.

Your cooperation is appreciated. Any questions should be referred to Gerald Nehls, NADB, MDAD, OAQPS."

The data bases of primary interest in the EPA Users Survey are:

NEDS

SAROAD

SIP Rules and Regulations

QAMIS (Quality Assurance Management Information System)

SOTDAT (Source Test Data System)

HATREMS (Hazardous and Trace Emissions Management System)

Air Quality Models

Many of these systems and data bases are already operational. Others are planned or are currently being developed. The following pages are brief synopses of each system.
Questions of concern are included, and personnel will be asked to respond to these questions, in addition to providing their own comments on current systems. The contractor's representatives will be prepared to discuss each system in more detail.

SYSTEM: NEDS

ACCESS: BATCH, REMOTE BATCH

OUTPUTS: Reports

POINT SOURCE LISTING	B,RB
AREA SOURCE LISTING	B,RB
AREA SOURCE FUEL USAGE	B,RB
COUNT OF PLANTS, POINTS, SCC's	B,RB
CONDENSED POINT SOURCE LISTING	B,RB
ALLOWED vs COMPUTED EMISSIONS	B,RB
EMISSION SUMMARY	B,RB
STATIONARY SOURCE FUEL SUMMARY	B,RB
SCC EMISSIONS REPORT	B
HIGHEST PLANT BY COUNTY	B
PLANT EMISSIONS SUMMARY	B
MISSING DATA ITEMS	B
WEIGHTED SENSITIVITY ANALYSIS	B
GRIDDING CAPABILITY	B
INPUT TO MODELS	B
DATA TAPE	B

SIGNIFICANT COMPUTERIZED ASPECTS:

Use of emission factors to calculate emissions, freedom to have various definitions of point sources without limiting nationwide comparisons and analyses, flexibility in SCC coding, remote batch access capability for scientists and engineers (not just computer specialists)

CURRENT DEVELOPMENT ACTIVITY:

The following capabilities are part of current plans for NEDS changes:

1. The TSO capability is being expanded to include more programs and to add selection criteria for limited output.
2. Programs for retrieval according to estimation method are planned.
3. A program will be added to produce a report of calculated potential emissions for a source.
4. A capability will be added to determine the effects of NEDS parameter changes on total emissions. (e.g. - what would be the result if all sources of a specified size burned 3% sulfur fuel.)
5. The area source input form is being reformatted. The primary change will be in Card 5.
6. If AQMA's follow county boundaries, a capability can be added to allow retrieval by AQMA.
7. A lowspeed terminal limited update capability will be available to the Regional Offices.
8. Regional Offices will be given the ability to run NADB edit programs from their terminals.
9. Validation routines will be available. (e.g. look at all sources >100 TPY but without controls.)

DEVELOPMENT CONSIDERATIONS:

Following are some considerations for future development of the NEDS.

1. Is CRT (display tube) terminal access to NEDS desireable? What applications do you have for CRT capability? How often would a CRT be accessed?
2. Would a gridding capability be desireable for application to air quality models? What might the gridding requirements be?
3. What additional data elements would be desireable in NEDS?
4. Are there elements in NEDS data that are not used?
5. Which currently available reports are used most often?
6. Are changes desireable for any of the retrieval options (i.e. batch, remote batch, or interactive)?
7. What suggestions do you have for updating procedures?
8. What kinds of management (non-technical) reports would be useful?
9. What long-range changes to NEDS are desireable? (i.e. after 1975.)
10. If a general statistical package is developed for NEDS/SAROAD, what statistics would be desireable for NEDS data?

SYSTEMS: SAROAD

ACCESS: BATCH, REMOTE BATCH, INTERACTIVE

OUTPUTS: Reports

QUARTERLY INVENTORY	I
QUARTERLY SUMMARY STATISTICS	I
YEARLY INVENTORY	I
YEARLY SUMMARY STATISTICS	I
UNITS CODES/NAMES TABLE	I
POLLUTANT CODE/NAME	I
QUARTERLY FREQUENCY DISTRIBUTIONS	B,RB
YEARLY FREQUENCY DISTRIBUTIONS	B,RB
RAW DATA LISTING	B,RB
STANDARDS REPORTS	B,RB
INVENTORIES (By Site, Pollutant, or Pollutant Within State)	B,RB
YEARLY REPORT BY QUARTERS	B
COMPOSITE LISTING	B
STATUS REPORTS	B
QUARTERLY REPORTS	B
PARAMETER/METHOD INFORMATION LISTING	B
DATA TAPE	B

SIGNIFICANT COMPUTERIZED ASPECTS:

Data editing, criteria for summarizations of data, criteria for incorporation of data, ability to provide some results in "Standard" units and other results in reporting units, interactive and remote batch capability for scientists and engineers (not only computer specialists)

CURRENT ACTIVITY:

Current plans for additions and changes to SAROAD include the following.

1. TSO capability will be expanded to add reports and limited selection criteria.
2. Standards printouts will be updated each time a data update is made.
3. Reports will be available for a sampling site only when data exists for the specified time period.
4. SAROAD reports by AQMA will be available if AQMA's follow county boundaries.
5. Regional Offices will have lowspeed terminal limited update capability.
6. Regional Offices will be able to run NADB edit programs from their terminals.
7. Regional Offices will have access to data validation programs. (e.g. high value checks, etc.)
8. A capability will be added to identify a set of coordinates and find all sampling sites within those coordinates.

DEVELOPMENT CONSIDERATIONS:

Following are some considerations for future development of SAROAD

1. Is CRT (display tube) terminal access to NEDS desireable? What applications do you have for CRT capability? How often would a CPT be accessed?
2. Which currently available reports are used most often?
3. Are changes desireable for any of the retrieval options (i.e. batch, remote batch, or interactive)?
4. What suggestions do you have for updating procedures?
5. What are your major concerns about SAROAD data validity? What are your suggestions for validation routines?
6. What kinds of management reports would be useful?

7. Are graphics desireable for displaying statistical summary data? What kinds of graphics? What statistics? How would graphics be used, i.e. how would this capability save time or improve on summary data already available?

8. What mechanism could be applied to track missing SAROAD data on a timely basis?

9. Are there reports not listed above that would be useful? What are the applications?

10. What capabilities should SAROAD have to be more useful in conjunction with other systems?

11. What long-range changes to SAROAD are desireable? (i.e. after 1975)

12. If a general statistical package is developed for NEDS/SAROAD, what statistics are desireable for SAROAD?

SYSTEM: SOTDAT (Source Testing Data System)

ACCESS: BATCH

OUTPUT: Reports

SOURCE DATA EMISSION FACTORS
PLANT CONTROL EQUIPMENT BY POLLUTANT
SOURCE TEST STATISTICS BY SCC

USE: Handle stack test data

CURRENT ACTIVITY: Installation on UNIVAC 1110 at RTP.

DEVELOPMENT CONSIDERATIONS:

Following are some considerations for future development of SOTDAT

1. What applications do you have for stack test data?
2. How is stack test data currently filed and retrieved?
3. What kinds of reports of stack test data would be useful? Why would these reports be useful?
4. What do the States need that might be provided through this system?

SYSTEM: QAMIS (Quality Assurance Management Information System)

ACCESS: BATCH

OUTPUTS (INITIAL).

There are at present five suggested reports to be developed by the contractor and approved by EPA:

1. site information
2. site-pollutant information
3. laboratory information
4. agency information
5. comprehensive of all of the above

Retrieval of quality assurance information will fall into one of these five categories.

USE: Handle Quality Assurance Information

CURRENT ACTIVITY:

Under development. Questionnaires on state and Regional Office quality assurance procedures for air quality monitoring stations have been completed. These will be computerized, and they will form the data base for the system.

DEVELOPMENT CONSIDERATIONS:

1. How can you apply this information to current procedures?
2. How can this system be responsive to your needs for tracking and validating SAROAD data?
3. Do you need the capability for storage and retrieval of estimates of precision for air quality data?
4. This is an interim system, and the base information will probably not change. What are the system requirements for long-range needs for quality assurance data?

SYSTEM: HATREMS (Hazardous and Trace Emissions Subsystems)

ACCESS: BATCH

OUTPUTS (INITIAL):

Not fully defined the equivalent of:

AREA SOURCE FUEL USAGE
CONDENSED POINT SOURCE LISTING
STATIONARY SOURCE FUEL SUMMARY
COMPUTED vs ALLOWED EMISSIONS

USE: Develop system to be used in conjunction with NEDS to handle up to 34 different pollutants

CURRENT ACTIVITY:

Under development. Projected completion for mid-1975.

DEVELOPMENT CONSIDERATIONS:

1. How do you currently handle hazardous and trace emissions data?
2. What access capability is required for these data?
3. What other reports and retrieval options would be useful? How would they be applied?
4. What features could make HATREMS more applicable to enforcement activities?
5. What are the immediate and long-range anticipated needs for hazardous emissions data?

SYSTEM: SIP Rules and Regulations

ACCESS: BATCH

OUTPUT: Rules and regulations, current or superceded, from State Implementation Plans. No Federal Regulations or Local Regulations, except for Local Regulations that are part of the SIP. Retrieval by geographic codes. Retrieval by criteria codes similar to SCC.

CURRENT ACTIVITY: Currently under development. Projected availability is Sept. - Oct. 1974.

DEVELOPMENT CONSIDERATIONS:

1. Would access other than batch be useful? What would the access be?
2. Would limited CRT (display tube) terminal access be desireable? What would the applications be? How often would the system be accessed?
3. What retrieval and report options are desired? How would these improve current procedures?

SYSTEM SUMMARY

At the beginning of each interview in the Regional Offices, the following summary of proposed systems changes was presented.

NEDS

1. Programs are planned which will allow for emission factors investigation, i.e. a listing of the number of plants by SCC and by a specific method of estimation.

2. Area source form will be reformatted. The change will involve Card 5 primarily. Details of the change will be distributed by NADB.

3. The edit check programs will be provided to the Regional Offices to facilitate data input.

4. TSO will be expanded to add programs and to add selection criteria: Regional Offices will have the same selection capability as NADB.

5. A report on potential emissions is planned. Potential emissions will probably be defined as emissions calculated by using maximum design capacity and/or assuming no control devices.

6. Provision of a statistical package is being considered and will be made available if there is sufficient need by the Regional Offices.

7. A capability may be added for selective retrievals based on AQMA's. This currently is planned only if the AQMA's follow county boundaries.

8. A system will be developed to allow analysis of potential changes to NEDS data. This capability will be useful for control strategy testing.

9. All programs written by anyone using the system should be documented according to NADB standards. NADB will maintain these programs. (i.e. Changes to configuration and file descriptions necessitated by hardware and software changes will be made by NADB).

10. User programs not documented according to NADB standards will not be maintained by NADB. A subfile of such programs for each Regional Office will be maintained, however, procedures for such maintenance have not as yet been finalized.

11. A lowspeed terminal limited update capability will be provided to the R.O.'s. This would bypass the magnetic tape requirement. Initially, some limit such as 50 cards would be placed on this capability.

12. Mobile source emission factors by county will be added. This information will be compiled from the Polk tapes. This information is currently handled on a state basis. The data will be used for area source calculations.

13. R.O.'s will have the capability for running NADB edit routines from terminals. NEDS edit is projected for March '75. Output would be routed to the R.O. terminal.

14. Validation routines will be available to the R.O.'s. These routines will point out suspicious data once they are on file, e.g. NEDS plants with stack heights >900 meters. Plans currently do not include checking specific problems for multiple criteria pollutants.

15. A program will be added to identify a set of coordinates and to find all plants within the coordinates.

16. Procedures are currently in progress in NADB to identify missing large point sources. Feedback will be to the R.O.'s, then to the States. 20 K missing points have been identified.

SAROAD

1. Standards printouts will be updated each time a data update is made. Currently the printout is updated quarterly only.

2. TSO capability will be changed to add programs and to add selection criteria: Regional Offices will have the same selection capability as NADB.

3. Reports for sampling sites will be available only when data exists for a specified period. The current system does not differentiate between active and discontinued sites.

4. Statistical packages will be added if needed.

5. Reports by AQMA will be available if AQMA's follow county boundaries.

6. The same documentation requirements as for NEDS programs written by the R.O.'s will apply to SAROAD programs written by the R.O.'s.

7. A lowspeed terminal limited update capability is planned. This will bypass the magnetic tape requirement. The update will initally be limited to 50 cards.

8. R.O.'s will have the capabillty for running NADB edit programs from their terminals. Output will be routed to the R.O. terminal. SAROAD is projected for January '75.

9. R.O.'s will have validation routines. e.g. - high value checks, or compare a quarterly average with a previous quarter average or with the same quarter the previous year.

10. A capability will be added to find all sampling sites within an area specified by the coordinates of a polygon shaped area.

SIP Rules & Regulations

1. The system is currently under contract to the MITRE Corporation.

2. Being written in SYSTEM 2000 language.

3. The system will have geographic codes and criteria codes that are not as complete as SCC codes.

4. The system initially will contain no local regulations, except those submitted with the SIP's.

5. The system will contain no Federal Regs, initailly. They may be included later.

6. The system will cover only rules in force or suspended.

7. Gives the number of regulations or the full text for the rules selcted. The capability to retrieve by regulation number is not included.

8. Approximately 6000 rules and regulations are being entered in the system.

9. Another contract, possibly in early '75, will tie SIP's into NEDS. The capability will be added to get SIP rules by NEDS point or to get NEDS points covered by a SIP rule. A link between SIP and CDS will be provided also.

10. NEDS Card 5 is not related to this system.

11. The SIP system will have AQCR codes only if the state regulations follow AQCR boundaries. Otherwise, IBM codes will be used.

12. Procedures for loading and updating this system have not been finalized.

QAMIS - Quality Assurance Management Information System

1. The QAMIS is an interim system based on the DQIS questionnaire that was distributed.

2. The questionnaire was a one-time effort to try to define the quality of the NASN operations.

3. Further development of a quality assurance system depends on QAEML policy.

HATREMS - Hazardous and Trace Emission System

1. HATREMS will handle up to 34 pollutants.

2. HATREMS will have a link to NEDS. If a plant address, etc. is in NEDS, it will not be duplicated in HATREMS.

3. Reports will parallel NEDS to a large extent, but will be more limited. Specific reports have not been defined to us.

4. Program development and updating efforts for HATREMS will not parallel those for NEDS, i.e. when a NEDS program change or addition is made, a HATREMS change or addition will not necessarily be made.

5. Projected completion for HATREMS is mid-'75.

CONSOLIDATED APPLICATIONS COMMENTS

REGIONAL OFFICES

NEDS

1. NEDS seminars indicated
2. Input latitude - longitude with subsequent conversion to UTM.
3. Define ranges of parameters for retrieval
4. Expand the comments field
5. NEDS/CDS cross reference needed
6. Eliminate Verification Report
7. Include SCC's codes for fugitive dust
8. Develop a polygon-defined retrieval capability
9. Statistical package not needed, except for general accounting
10. Confidentiality reporting requirements cause problems with data retrieved.
11. Need a report to show total plant emissions
12. Hand calculated vs. machine calculated emissions sometimes confusing
13. Revisions to NEDS instruction manual (AP-42) are needed
14. Revisions to the NEDS User Manual are needed
15. Require identification of data source
16. Lack of positions at Regional Office and State levels cause problems
17. Most states and R.O.'s not interested in updating area source data

18. Reporting requirements should not be changed

19. EIS implementation should improve the data base

20. States lack incentives to update NEDS.

SAROAD

1. Interactive access needed for all short reports

2. Should have option to include data not meeting 75% criterion in data summaries

3. Include 2nd highest value in summary reports

4. Polygon area retrieval is desirable

5. A general statistical package is not needed

6. Frequency distribution for 8-hour CO and 3-hour and 24-hour SO_2 averages is needed

7. Plotting capabilities needed:
   high 8-hour CO ave./week/yr.
   High 24-hour SO_2 ave./week/yr.
   monthly averages
   yearly averages
   diurnal variations
   % of days/year with short-term violations

8. Bypass validation rejects if desired

9. Project codes cause problems

10. Tighten the minimum requirements for averaging Hi-vol data

11. Flag SIP required stations on reports

12. Audit data from SIP required stations

13. Identify inactive sites

14. Discontinue interactive quarterly and yearly inventory reports

15. Implement an interactive report identifying the period of most recent data for sites

16. Implement a capability to track site identification numbers. Changing numbers cause problems in historical data analysis.

17. Implement a report of values exceeding standards - include site code, county code, No. of violations of primary standards, % of values greater than primary standards

8. Show raw data on edit and validation reports in reporting units

19. Compare running averages against standards for plotting to make trends projections

20. Add an area code to the Status Report

21. Plot trends as new data become available

22. Select output by county and AQMA

23. Allow laboratory to input basic parameters and have the computer determine concentrations

24. Verification Report should list only new data - not all data. Limit volume of these kinds of reports

25. Prepare wind roses and pollutant roses

26. Validation should check data for a specific site for data anomalies

27. Option to interactively request multiple copies of batch reports

28. Only one R.O. requested CRT capability

29. Implement an audit of CPU and connect time by report type

30. Include pollutant/method description in site description

31. Prepare 12 month running averages

32. Indicate percent improvement in air quality for specific sites

33. Eliminate some of the conversational aspects of multiple batch requests

34. Include short-term episode monitoring data
35. Discontinue the validation report - rather flag questionable values during edit
36. Implement interactive capability to determine what is monitored at a site
37. Expanded SAROAD city codes to account for burroughs, cities, townships are needed
38. Implement SAROAD parameter listing as an interactive report
39. Correct data for altitude effects
40. Change "percent valid data" to "percent available data"
41. Oxidant method codes should indicate if correction for NO or SO_2 interference has been made
42. Retrieve data in a county for sites in unincorporated areas
43. Include the mean and standard deviation for 3-hour and 8-hour averages
44. Reports should have the same number of significant digits
45. Present the standard geometric deviation to 3 decimals
46. Implement a trends report for 12 month, 24 month, and 36 month moving averages

HATREMS

1. No significant applications in any R.O.'s for this system.

SOTDAT

1. No significant applications were identified.
2. Region I would like to use to see what other states have done.

SIP Rules and Regulations

1. Interactive capability is necessary to make this system meaningful.

2. System should include state approved as well as federally approved regulations.
3. A report of the most current update numbers for rules might be useful.
4. No immediate requirements for this system were identified.

QAMIS

1. It would be possible to assign quality evaluation numbers for each station, but several numbers would have to be assigned - one for each pollutant method at the station.
2. Quality assurance is best controlled by site visits and agency evaluations.
3. R.O.'s should have input before any quality assurance requirements are promulgated.
4. The judgment of data quality must be subjective for each site.
5. Data generators should tag data as good or bad rather than relying on a system to do it. Subjective judgment is required.
6. Assigning grades for data quality would be politically unwise and would jeopardize the cooperation of the states in implementing quality assurance programs.

CONSOLIDATED APPLICATIONS COMMENTS

EPA USERS OUTSIDE R.O.'s

NEDS

1. DSSE should have access to NEDS reports via terminal.
2. Estimation code to indicate "114" letters as data sources.
3. NEDS-CDS cross-reference needed - will help facilitate R-45 Report.
4. Expanded comments field needed.
5. Transmit update data to CDS when NEDS is updated.
6. Geographic Location Codes should replace SAROAD city/county codes.
7. One or two day report turnaround required.
8. NEDS user manual describing reports, retrieval options, etc.
9. User seminars (as opposed to contributor seminars) are needed.
10. Establish emissions trends.
11. If potential emissions are not considered, significant sources can be lost from the system.
12. One digit emission category.
13. Establish emissions trends by area and by source type.
14. Polygon retrieval.
15. Include a NEDS-FPC '67 cross-reference number.
16. Query capability to indicate the amount of data available.
17. Three to four day turnaround required.
18. Do not make NEDS reporting requirements less stringent.
19. Identify data source.
20. Retrieval by ranges of parameters.
21. Publish NADB information in the "EPA Systems News", published by MIDSD.

SAROAD

1. Polygon retrieval.
2. SIP station audit capability.
3. Monthly submittal from states should be a goal.
4. Retain the agency/project code. It is important for trends monitoring.
5. Expand the capabilities of the method code. Teach people how to use it.
6. Add an active/inactive code, dates, station cross-reference numbers, etc. to the Site Description File.
7. Retain the job status query capability.
8. Modify alignment of print wheel on the DCT 500.
9. Implement SAS.
10. Parametric data retrieval.
11. Plot seasonal trends.
12. Status report to show how much data, pollutants, years, etc. are available for a station.
13. User manual to allow requests in English rather than in codes.
14. Print a synopsis of available data on a map. This could be a publication.
15. Allow users to read data directly from RTP computer into WYLBUR file at OSI.
16. Table to describe retrieval options.
17. County retrieval.
18. Include minimum detectable level on reports.
19. Statement of method accuracy in the Site Identification File.
20. Identify SIP stations in Site Identification File.
21. Pollutant roses and meteorological roses.
22. Report of standards violations.

HATREMS

No significant input for HATREMS.

SIP Rules and Regulations

No significant input

SOTDAT

No significant input